Moral Judgment and Decision Making under Uncertainty

D i s s e r t a t i o n
zur Erlangung des akademischen Grades
Dr. rer. nat. (doctor rerum naturalium)
im Fach Psychologie

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät II
der Humboldt-Universität zu Berlin

von
Nadine Fleischhut, M.A. phil.

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:
Prof. Dr. Elmar Kulke

Gutachter:
1. Prof. Dr. Gerd Gigerenzer
2. Prof. Dr. Elke van der Meer
3. Prof. Dr. Werner Güth

Eingereicht am: 6. März 2013
Verteidigt am: 6. August 2013

Table of Contents

Chapter 1

Introduction

Uncertainty itself is neither ethical nor unethical—yet it is inherent to most situations in which moral judgments and decisions have to be made. Would you recommend breast cancer screening given a small benefit and a risk of wrong diagnosis? Should an airplane be shot down if there is a risk that it may have been kidnapped by terrorists? To what extent does each of us have to reduce her energy consumption to prevent severe climate change? And what are the benefits of donating to someone in need on the street if it is uncertain how that money will be spent? In many moral situations, neither the future course of events nor the consequences of our actions and omissions are known for sure. The uncertainty under which we live can have important implications for our moral judgments and decisions and, consequently, for the well-being of others.

In this dissertation, I investigate the question of how people make moral judgments and decisions under uncertainty. The three research projects, intended for publication as stand-alone articles, present theoretical and empirical work that shows the importance of studying judgment and decision making in moral situations under epistemic conditions characteristic of real-world situations. Theoretically, judgment and decision making in moral situations is analyzed from the perspective of bounded (Gigerenzer, Todd, & The ABC Research Group, 1999; Simon, 1979) and ecological rationality (Todd, Gigerenzer, & The ABC Research Group, 2012). The former emphasizes the cognitive and informational

constraints under which judgments and decisions have to be made, whereas the latter focuses on the interplay between cognition and the structure of the environment. Both notions thus imply the importance of considering the uncertainty under which we live, judge, and decide. Empirically, the goal is therefore to investigate moral judgment and decision making under the different imperfect epistemic conditions people encounter in the real world. The empirical part of the dissertation focuses on risk and uncertainty in two widely used research paradigms within the study of morality: cooperation decisions in social dilemmas (for a review, see von Dijk, Wit, Wilke, & Budescu, 2004) and judgments in moral dilemmas (for a review, see Waldmann, Nagel, & Wiegmann, 2012).

In this dissertation, the moral domain is not defined by a particular normative moral theory or restricted to the application of particular moral rules or norms (for a critique of restricting the moral domain by Western or rationalistic notions of morality, see Haidt & Joseph, 2007). Even more than in other domains, there is no agreement in ethics about the correct normative moral theory—a dispute that may well be principally undecidable.[1] Instead of arbitrarily subscribing to some version of consequentialism (e.g., Baron, 1994; Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008; Sunstein, 2005) or a Kantian framework (e.g., Kohlberg, 1984), the work presented here instead understands judgments and decisions in moral situations from a social-functionalist perspective of what they are good for: as norms, practices, and evolved psychological mechanisms that regulate social relationships and promote the coherence of social groups (Haidt & Kesebir, 2010; Rai & Fiske, 2011).

Before presenting the three research projects in Chapters 2 to 4, I will start by introducing the notions of bounded and ecological rationality, which provide the general theoretical framework for the research presented here. This will be followed by a short introduction regarding the practical and theoretical relevance of uncertainty for the study of judgment and decision making in moral situations. I will conclude with a brief overview of the three studies and the specific aspects they focus on.

[1] Normatively, "morality" is used to refer to a code of conduct that under certain conditions would be endorsed every rational person. What counts as immoral will thus differ depending on the normative theory proposed. Descriptively, morality refers to the codes of conduct adopted by groups or individuals which thus vary in their content across groups, times and cultures. It is an open question whether moral cognition nevertheless differs from other kinds of cognition, for instance, by relying on a special kind of norms (Kelly, Stich, Haley, Eng, & Fessler, 2007; Turiel, 1983), relying on domain-specific cognitive processes (e.g., Mikhail, 2007), or triggering particular emotions (e.g., Haidt, 2001; Hume, 1983).

Bounded and Ecological Rationality in the Moral Domain

The study of bounded rationality asks the question how human beings make judgments and decisions when information, time, or cognitive capacities are limited (Simon, 1979). Contrary to the requirements of many rational models, our knowledge is generally far from perfect. Following Herbert Simon, researchers in the tradition of bounded rationality have thus been concerned with understanding the heuristics people rely on under realistic epistemic conditions. Heuristics are cognitive processes that work efficiently by ignoring information (Gigerenzer & Gaissmaier, 2011)—and thus are particularly applicable cognitive tools when alternatives, consequences, or probabilities are unknown. Some heuristics ignore weights (e.g., Dawes, 1979; Einhorn & Hogarth, 1975) or probabilities (e.g., Wald, 1945). Others do not integrate the available information but instead consider cues sequentially until a decision can be made (Gigerenzer & Goldstein, 1996; Tversky, 1972). Different from general-purpose models, heuristics are specialized processes adapted to particular environments.

The study of ecological rationality, in turn, investigates in which environments a given heuristic will work and where it will fail (Todd et al., 2012). Contrary to the view that heuristics necessarily trade off time for accuracy (e.g., Beach & Mitchell, 1978; Payne, Bettman, & Johnson, 1993), heuristics can, under certain conditions, be even more accurate than complex models (Gigerenzer et al., 1999). This is achieved precisely by ignoring part of the information and exploiting the match between the process and the structure of the environment. Even if specialized heuristics are less flexible than general-purpose models, they can make better predictions by being less sensitive to variance in observations, which particularly pays off if the experienced samples are small (Gigerenzer & Brighton, 2009). Within the study of ecological rationality, the epistemic conditions that people face are an important part of the structure of the environment without which cognitive processes cannot be understood. The framework of bounded and ecological rationality thus emphasizes the uncertainty under which we live, and which in turn shapes the cognitive processes underlying our judgments and decisions.

While the approach has been successful in explaining judgment and decision making in many domains (Gigerenzer et al., 1999; Payne et al., 1993), it has not yet received the same attention within the study of morality (but see, e.g., Chugh, Bazerman, & Banaji, 2005, for a notion of bounded ethicality). Note that linking the study of morality to the

framework of bounded and ecological rationality does not mean equating the norms of morality and rationality. To evaluate something on a moral dimension is not the same as evaluating its rationality. Yet for the descriptive understanding of judgment and decisions in moral situations, it is an important lesson to acknowledge both, the cognitive side and the environment—and thus the uncertainty of the world and how the mind deals with it.

Uncertainties in Moral Judgment and Decision Making

Moral situations allow for different sources of uncertainty, which can be *epistemic*, *social*, or *normative*. *Epistemic uncertainty* refers to the limited information and predictability of the natural world. Judgments and decisions in moral situations often have to be made when the outcomes of our actions are not certain. Yet insofar as such judgments and decisions depend on non-normative information about the situation, uncertainty and limited predictability are not without consequences. The epistemic situation people face can influence how a situation is perceived, affect the moral or non-moral decision processes that are applied, or even restrict the decision rules that are applicable depending on the information that they require.

According to a seminal classification by Knight (1921), three types of uncertainty can be distinguished. In some situations, we can calculate the objective probabilities of events a priori, for instance, the chances of a (fair) die showing a six. In other cases, probabilities can be estimated statistically from data, for instance, the chances of being wrongly diagnosed of cancer when participating in regular screenings. Both cases are referred to as situations of *risk*. Finally, there are situations with so many unique features that they can hardly be grouped with similar cases, such as the danger resulting from a new type of virus, or the consequences of military intervention in conflict areas. These represent cases of (Knightian) uncertainty where no data are available to estimate objective probabilities. While we may rely on our subjective estimates under such conditions, no objective basis exists by which to judge them (e.g., LeRoy & Singell, 1987).[2] However, even when information is in principle attainable and chances could be estimated, an individual agent or

[2] The interpretation of Knight's distinction is under dispute. It has been understood as a distinction between what is measurable vs. unmeasurable, objective vs. subjective probabilities (Friedman, 1976), or insurability vs. uninsurability (e.g., LeRoy & Singell, 1987). Depending on the interpretation, the distinguishing features of uncertainty differ, as does the explanation for Knight's claim that for estimates under uncertainty there is no valid basis for classifying instances (see, e.g., Langlois & Cosgel, 1993).

agency may still lack the time or resources to do so—and thus face a situation of uncertainty.

The range that goes from precisely calculable a priori probabilities to situations of uncertainty where no similar cases are available can also be understood as a continuum of *degrees of uncertainty*, that is, as a function of the number of similar past experiences from which probabilities can be inferred (e.g., Rakow & Newell, 2010).[3] Corresponding to these degrees, different external representations can be distinguished, from complete summary descriptions in terms of probabilities, which are typically used to convey risk information, to the sequential experience of larger or smaller samples common to most daily life situations. Importantly, studies on risky decision making have found a pronounced difference in choices depending on whether risk information is learned through a *description* of probabilities or through *experience* (labeled the Description–Experience (DE) gap; Hertwig, Barron, Weber, & Erev, 2004). Yet the different degrees of uncertainty and the way information is acquired have not received enough attention in the study of moral situations (for reviews see, e.g., Christensen & Gomila, 2012; von Dijk et al., 2004).

The problem of *epistemic uncertainty* needs to be distinguished from *social uncertainty* about the action of others. *Social uncertainty* represents a second major source of uncertainty relevant to many moral situations, as these are typically situations of social interdependence. In moral situations, others depend on us as we often depend on them, even when they are not directly present. For instance, in common social dilemma situations, it is in one's self-interest not to cooperate, while it would be best for the group if everyone cooperated. What we gain from cooperation thus depends on the behavior of others. While social uncertainty is not the main focus of this dissertation, the perspective of ecological rationality emphasizes the importance of studying moral judgments and decisions in the social environments in which they are typically made.

Finally, people may face *normative* or *moral uncertainty* about which norm to follow, and how to act in the face of conflicting or ambiguous moral theories. Even with perfect knowledge about the world and about the behavior of others, such *normative uncertainty* would still remain. The issue of moral uncertainty is a normative question and that is outside the scope of this dissertation. Yet it is important to note that empirical results have implications for such normative questions. Normative uncertainty may arise

[3] See Brighton and Gigerenzer (2012) and Lo and Müller (2010) for a more refined taxonomy that distinguishes qualitatively different categories of uncertainty based on their source.

when the implications of a moral theory for a given situation are unclear, when the situation is ambiguous, or when information is missing. For instance, theories such as *consequentialism* (for an introduction, see Scheffler, 1988) require as much information about consequences, alternatives, and probabilities as do models of rational expectations (Gigerenzer, 2010), and may thus neither be able to provide feasible moral recommendations, nor be good candidates for peoples' decision process (Bennis, Medin, & Bartels, 2010; Binmore, 2008; Gigerenzer, 2010). Ignoring epistemic uncertainty in the formulation of normative theories can thus result in practical uncertainty about what to do in a given situation (Sollie, 2009).

In summary, the practical importance of *epistemic uncertainty* for the study of moral judgment and decision making results from the risks and uncertainties that pervade everyday situations, but also from acknowledging the challenges that arise in cases of real Knightian uncertainty. Theoretically, uncertainty is important for the study of moral judgment and decision making because the heuristics people may use in moral situations may be particularly adapted for judgments and decisions under limited information. Furthermore, these heuristics can also be efficient devices to reduce uncertainty, for instance, by relying on imitation or social learning processes. At the same time, they can provide orientation about the expected behavior of others and enhance the social coherence of the group. Thus, understanding the function of the heuristics people may follow in moral situations requires them to be studied under the epistemic conditions to which they may be adapted.

A Research Gap in Standard Paradigms of Moral Judgment and Decision Making

The practical and theoretical relevance of uncertainty for moral judgment and decision making does not correspond with standard paradigms used in the empirical study of morality. In this dissertation, I focus on two widely used empirical paradigms of moral judgments and decision making: the study of moral judgments in dilemma situations, such as the "trolley" cases, and the study of cooperation decisions in social dilemmas.

Moral psychology has long been a branch of developmental psychology due to the dominant paradigm of Kohlberg (1984). Within the last decade, however, the study of morality and, in particular, the study of moral judgment has attracted renewed and highly interdisciplinary interest (Haidt, 2007). What is new is first and foremost a move away from the rationalist paradigm of Kohlberg towards the study of intuitive moral judgments. Here,

moral intuitions refer to fast and often unconscious processes that bring about a (typically affect-laden) evaluation (e.g., Haidt & Bjorklund, 2008). While Kohlberg's rationalistic paradigm endorsed the implicit assumption that moral judgments have to be accompanied by reason, the new insight was that these reasons may often not be the causes of moral judgments but rather post hoc justifications (Haidt, 2001). In addition, people are not necessarily aware of the processes underlying their moral judgments (Hauser, Cushman, Young, Kang-Xing Jin, & Mikhail, 2007).

The shift towards the study of moral intuitions and the underlying factors and processes promises to be valuable for future research (Waldmann et al., 2012). Yet the use of ideal-typical moral dilemmas as experimental paradigms in the study of moral judgment has remained. Some of the most famous new insights (e.g., Greene, Sommerville, Nystrom, Darley, & Cohen, 2001) have been gained by using artificial moral dilemmas, some of which had previously been used as philosophical thought experiments. Typically, these dilemmas present the consequences of all options as certain and are thus not representative of the different degrees of uncertainty that people encounter in the real world. Yet from the perspective of bounded and ecological rationality, ignoring the epistemic conditions under which judgments have to be made will necessarily obstruct our understanding of moral judgment.

Due to the rationalist heritage and the assumption that moral judgments precede moral decision making (Kohlberg, Levine, & Hewer, 1983), the focus of moral psychology has been on moral judgment rather than on decision making (for the same assumption in business ethics, see Jones, 1991; Rest, 1986; Trevino, 1986). The study of moral behavior and decision making has instead been located within social psychology and behavioral economics, especially in research on altruism, pro-social behavior, and social preferences. Here, cooperation has posed a puzzle that has been investigated in a variety of research paradigms, with the study of social dilemmas and the "tragedy of the commons" being one of the most famous (Hardin, 1968; Vollan & Ostrom, 2010). Laboratory studies have typically included social uncertainty and considered epistemic uncertainty from the natural environment (often called "environmental" uncertainty) to some extent as well (Von Dijk et al., 2004). However, the different degrees of risk and uncertainty, and the way in which information is acquired—by a *description* of probabilities or by *experience* (Hertwig et al., 2004)—has not been taken into account.

In what follows, I will give a brief overview of the three studies presented in this dissertation and the different issues of moral judgment and decision making under uncertainty on which they focus.

Beyond Character: How to Explain Moral Inconsistencies? This dissertation starts in Chapter 2 with an analysis of judgment and decision making in moral situations from the perspective of bounded and ecological rationality. (Un)ethical behavior is typically explained by characteristics of the individual, such as character traits (Foot, 1978), social preferences (Fehr & Schmidt, 1999), or moral stages (Kohlberg, 1984). These internalistic explanations predict ethical behavior to be stable across a range of situations. To account for commonly observed moral inconsistencies, these theories often resort to ad hoc explanations—referring, for instance, to cognitive biases (e.g., Messick & Bazerman, 2001), situational moderators (e.g., Trevino, 1986), or a temporary blindness to the ethical aspect of a decision (for a review, see Tenbrunsel & Smith-Crowe, 2008).

In contrast, we discuss how behavior that appears inconsistent from an internal point of view can be consistent when the interaction between heuristic processes and the environment is taken into consideration. Adopting the perspective of ecological rationality can thus help to explain three types of moral inconsistencies that are commonly observed in moral judgments and decisions.

Second, we argue that much judgment and behavior in moral situations under uncertainty may result not from specifically moral rules but from morally neutral social heuristics. Not only do moralities differ across time and culture (Haidt & Joseph, 2007; Rozin, 1999; Shweder, Much, Mahaptra, & Park, 1997), but the same heuristics can be used in moral and non-moral situations. Instead of defining the moral domain based on a biased Western notion or arbitrary normative theory, we adopt a functionalist perspective of morality as a means to regulate social relationships (Haidt, 2007; Rai & Fiske, 2011). The perspective of ecological rationality further emphasizes the need to study epistemic and social environments in order to understand judgment and behavior in moral situations. By using social heuristics, such as imitation strategies (Axelrod, 1986), the equity heuristic (Hertwig, Davis, & Sulloway, 2002; Messick & Schell, 1992), or defaults (Johnson & Goldstein, 2003), people may be able to reduce the uncertainty that they face, while at the same time preserving the social coherence of the group.

Decision Making in Social Dilemmas: Cooperation in Risky Environments When people cooperate, the results of their efforts are typically uncertain. On the one hand, joint efforts often pose a social dilemma where it is in the interest of all to cooperate; yet individually it is better to free-ride on the efforts of others. Social dilemmas thus include *social uncertainty* about the action of others. On the other hand, even if everyone cooperates, the outcomes of cooperation are often uncertain due to risk in the environment.

Chapter 3 empirically examines how cooperative behavior is shaped by different levels of risk and by the way information about risk is acquired (from *description* or from *experience*). In experimental social dilemmas, participants typically learn about the risky environment from a description of outcomes and probabilities (e.g., Gangadharan & Nemes, 2009; Levati, Morone, & Fiore, 2009) rather than by experience, without acknowledging that the two types of learning may call upon different decision processes which lead to different choices. In fact, research on risky choice showed that people's choices differ strongly depending on whether risk information is described in terms of probabilities or acquired through experience (Hertwig et al., 2004; Rakow & Newell, 2010).

We compared cooperation decisions in public good dilemmas based on a description of risk information with decisions based on learning by experience. In the *description* condition, participants received information about environmental risks as a probability statement, whereas in the *experience* condition, participants sampled to infer the probabilities from observed frequencies. To investigate how different levels of risk affect behavior, we varied probabilities and outcomes of the public good within subjects while keeping the expected outcome constant. Finally, we compared decisions in social dilemmas to those made in nonsocial lottery situations with identical levels of risk, to control whether the values and probabilities chosen to implement environmental risk replicate findings from risky choice in a standard setting.

Cooperation varied systematically with different levels of risk, mimicking behavior in nonsocial lottery situations. Whether people cooperate thus depends on the riskiness of the environment rather than on stable social preferences. Nevertheless, the way in which information was acquired—by description or by experience—mattered only for lotteries and surprisingly not for social dilemmas. Thus, no DE gap (i.e., difference in choices due to the presentation format) was found for decision making in social dilemma situations. Process data and self-reports suggest that this discrepancy between nonsocial lotteries and

social dilemmas may have resulted from decision processes that are more sensitive to beliefs about others' behavior and the size of rewards than to reward probabilities.

Moral Hindsight: Moral Judgments under Certainty and Uncertainty Uncertainty is inherent in many situations in which moral judgments are made, whereas the course of events is certain only after the fact. Chapter 4 of this dissertation examines judgments in moral dilemmas under uncertainty, as opposed to commonly investigated artificial dilemma situations, such as the "trolley" cases (e.g., Foot, 1967; Greene et al., 2001; Hauser et al., 2007), in which the course of all future events is described as certain. In particular, we compare judgments of moral permissibility in two types of epistemic situations: situations in which the course of events is still uncertain (i.e., when judgments are made in *foresight*) and situations in which it is certain whether negative side-effects did or did not occur (i.e., when judgments are made in *hindsight*). To examine how moral judgments are influenced by what is foreseen, we also asked participants to estimate the probability of side effects.

The key finding was a hindsight effect (Fischhoff, 1975; Hawkins & Hastie, 1990) for moral judgment. Actions were judged to be more morally permissible in foresight than in hindsight when participants knew that negative side effects occurred. Conversely, when participants knew that no negative side effects occurred, they judged actions to be more permissible than did foresight participants. As moral judgments differed under certainty and uncertainty, this raises concerns about the generalizability of empirical results from commonly investigated moral dilemmas in which everything is presented as certain—and which thus do not match the epistemic conditions people face in the real world.

A hindsight effect was also found for participants' probability estimates of side effects. Specifically, people overestimated in hindsight what they would have foreseen at the time when the decision was made, which again highlights the importance of considering the epistemic conditions under which moral judgments are made. The patterns of moral judgments and probability estimates align in a way that is consistent with a consequentialist process of moral judgment, according to which people weigh the possible consequences by the probability estimates adjusted in the direction of already known outcomes. However, probability estimates and moral judgments were only moderately correlated. In fact, an exploratory analysis of people's most important reasons for their judgments provided converging evidence that not everyone took probabilities about negative side effects into ac-

count. Because not all judgment processes may be similarly sensitive to uncertainty and missing information, this further underlines the importance of studying both the cognitive processes and the epistemic conditions to which they may be adapted.

Chapter 2

Can Simple Heuristics Explain Moral Inconsistencies?[1]

Nadine Fleischhut [a] & Gerd Gigerenzer [a]

[a] Max Planck Institute for Human Development, Berlin

[1] Fleischhut, N., & Gigerenzer, G. (2013). Can simple heuristics explain moral inconsistencies? In Hertwig, R., Hoffrage, U., & the ABC Research Group. *Simple heuristics in a social world* (pp. 459-486). New York: Oxford University Press.

Abstract

From virtue theory to moral psychology to behavioral economics, a range of disciplines have explained behavior in moral situations by states of the individual mind, such as character traits, moral stages, or social preferences. These internal explanations predict that moral behavior is stable across a range of situations and thus struggle with the common observation of inconsistencies in moral judgment and behavior. In contrast, we first outline how the same heuristic predicts systematically different outcomes, ethical or unethical, depending on differences in the environment. Behavior that appears inconsistent from an internal point of view is actually consistent when the interaction between heuristics and social environments is taken in consideration. Second, we argue that the heuristics determining much of judgment and behavior in moral situations are not specifically moral rules, but morally neutral heuristics that serve particular social goals. Specifying these processes can facilitate understanding when and why humans succeed or fail in pursuing ethical goals. The approach thus emphasizes the relevance of designing better environments, not just better people, in order to effectively promote the ethical goals valued by a society.

Keywords: Moral judgment and behavior, heuristics, social rationality, ethical decision making, adaptive toolbox, moral psychology, virtues

Introduction

Equality has been and continues to be one of the paradigmatic moral concerns for many Western societies. Who is to be treated as "equal" has been continuously redefined in an attempt to create legal institutions and practices that do not systematically discriminate against any group. In 1955, Rosa Parks, an African-American, was arrested in the United States when she refused to give up her seat to make room for a white passenger on the bus. And not until 1971 did women in Switzerland gain the right to vote in federal elections.[2] Yet concerns for equality are by no means restricted to the realm of legal regulation. Although many Western parents say that they aim to treat their children equally (Daniels, Dunn, Furstenberg, & Plomin, 1985; Schooler, 1972), studies nevertheless report a variety of birth-order effects. For instance, some children receive more care time from their parents than their siblings do. As Figure 1 shows, the distribution of care time within families varies as a function of the number of children. While each child receives equal time in families with two children, inequality in time allocation emerges in families of three, four, or five children. Furthermore, unequal treatment amplifies when the interbirth interval becomes larger.

But why would parents treat their children unequally? Do they favor particular children, consciously or unconsciously? Or do some children demand more time, while others are already seen as more self-reliant? Explanations along these lines would be different for particular children and parents, but are unlikely to produce the systematic pattern shown in Figure 1. Rather than focusing on idiosyncratic features, Hertwig, Davis, and Sulloway (2002) offered a more parsimonious account to explain a variety of systematic birth-order effects. They suggested that parents might rely on an equity *heuristic*[3]:

Allocate the resource equally to each of N alternatives (in this case, children).

Such a decision strategy is simple because it does not require much information processing. At the same time, its policy conforms to egalitarian values. Consider a family with two children, where the parents divide their care time equally per day or week according to the *equity heuristic*. As a result, the total care time each child receives during childhood will be the same, as shown in Figure 1 by the black squares. Now imagine that the

[2] The last Swiss canton to grant women the right to vote did so in 1990.
[3] Hertwig et al. (2002 used the term *equity*, whereas Messick & Schell (1992) referred to an *equality heuristic* for the same allocation principle. In research on distributional justice, *equity* usually means an allocation proportional to some criterion; that is, effort or need. If the criterion values do not differ, the equity principle amounts to equal allocation.

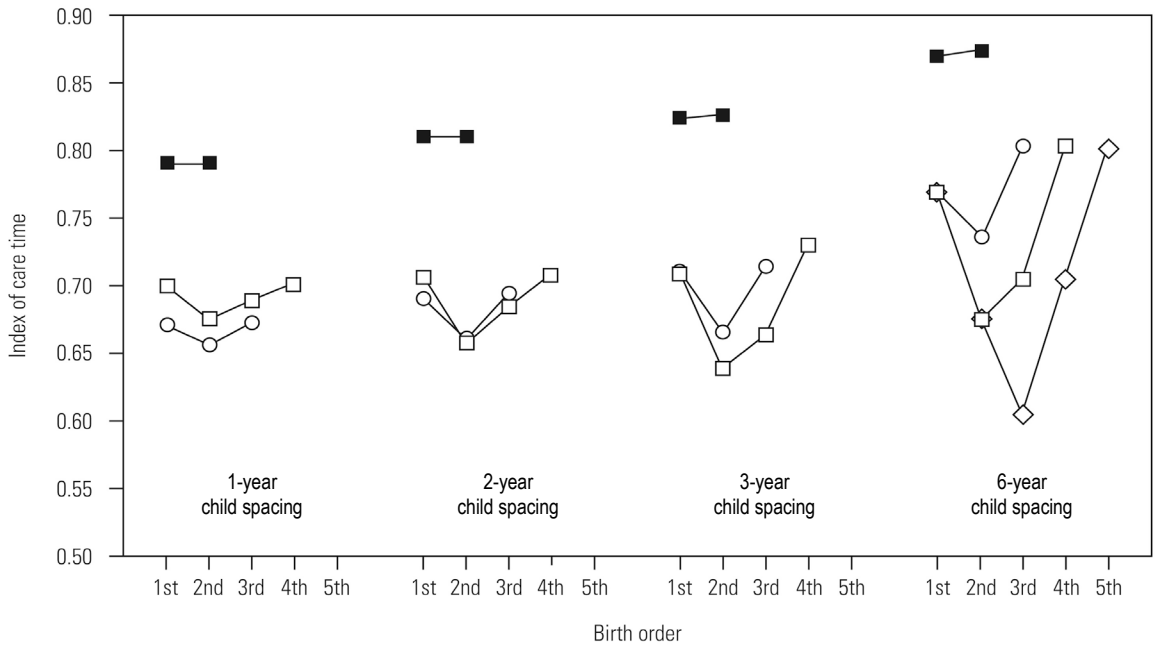


Figure 1. Distribution of total child-care time (y-axis) for families with different average birth intervals between children (x-axis). The distributions vary as a function of family size (with the number of children represented by the number of identical symbols) and birth rank (Source: Adapted from Figure 4 in "Parental investment: How an equity motive can produce inequality" by R. Hertwig, J. N. Davis, & F. J. Sulloway, *Pyschological Bulletin, 128*, 728–745. Copyright 2002 by the American Psychological Association.)

parents have a third child and continue to apply the same distribution strategy. The *equity heuristic* predicts that treating every child equally within each day or week will now produce unwanted birth-order effects in the long run: because middle-born children never experience a time when resources do not have to be shared, they receive, in the end, a smaller share of the total resources than their earlier- and later-born siblings.

The case of parental investment illustrates how the same heuristic can produce both an equal and an unequal distribution of resources—and thus a more- or less-just outcome—*depending on the environment*; here, the number of children and the duration of the interbirth intervals (see Figure 1). It is an example of how the interaction between a heuristic and the environment explains outcomes in a moral situation (Gigerenzer, 2010). We will use the term *moral situations* as shorthand for situations that concern moral issues; that is, for situations where behavior is frequently evaluated as moral or immoral according

to the standards of the individual himself, his or her peer group, or an even wider community (Haidt & Joseph, 2004; Shweder, Much, Mahaptra, & Park, 1997). Although moral standards may differ across social groups, many Western societies will consider distribution of resources to be a moral issue, and—in the case of parental investment—parents themselves may judge an unequal distribution within their family as morally problematic. Understanding parents' behavior as an interaction between a heuristic process and the environment thus offers one explanation of why and when they fall short of their own moral values.

This account can be contrasted with those that rely exclusively on internal explanations, such as character, moral rules, or social preferences. The theory of *inequity aversion* (Fehr & Schmidt, 1999), for instance, has been proposed to account for equality motives by means of a modified utility theory. Standard expected utility theory models behavior as a function of the utility of an action's possible consequences, which are multiplied by their probability and then summed up. An individual is predicted to choose rationally; that is, to choose the alternative with the highest expected utility. *Inequity aversion* adds two additional parameters to the standard theory, which capture how much an individual cares for equality (and despises favorable and unfavorable inequality). Thus, the theory explains behavior by focusing on stable preferences—an explanatory entity residing entirely within a person.

The view proposed in this chapter does not bet on internal explanations, but follows Herbert Simon's (1956) understanding of bounded rationality in analyzing the *match* of simple heuristics with the environments in which they succeed or fail. This facilitates understanding how people make decisions under uncertainty, with limited time and cognitive resources—and without utilities or probabilities (Gigerenzer, 2008; Gigerenzer & Selten, 2001). The case of parental investment demonstrates how the same cognitive process can result in either fair or unfair divisions, depending on the environment in which it is applied. It also shows that complex behavioral patterns, as shown in Figure 1, do not require complex explanations such as the calculations assumed in expected utility theory and its modifications. Given that behavior in moral situations is a form of decision making under uncertainty, we suggest that the framework of the fast-and-frugal heuristics program can be valuable to understanding such behavior as well (Gigerenzer, 2008; Gigerenzer, 2010; Marewski & Krol, 2010).

We propose two hypotheses, the first of which follows directly from the fast-and-frugal heuristics program:

(1) Judgment and behavior in moral situations may often (but not always) be the product of the interaction of heuristics and the environment, rather than of internal causes alone.

Unlike explanations that cite internal causes, such as virtues, character traits, or moral preferences, the focus on the match between environments and heuristics can, for example, help explain why inequality may prevail even under conditions in which people do act according to good intentions, such as when parents aim to treat their children equally.

The second hypothesis specifies the kind of heuristics that we believe play an important role in moral situations:

(2) Instead of resulting from the application of specific moral rules, judgment and behavior in moral situations may often be due to morally neutral simple heuristics applied in the social domain.

We focus on simple heuristics that are used in social games (Hertwig & Herzog, 2009) in which other people provide the most important aspect of the task environment—as in many moral situations. Some of these heuristics can also be applied in games against nature. One example is the *equity heuristic*, which underlies nonsocial decisions like how to allocate money across different investment options (DeMiguel, Garlappi, & Uppal, 2009), as well as social decisions such as parental investment. Furthermore, there are also heuristics that are more specifically "social," for instance *imitate-the-majority* or *tit-for-tat*: by processing social information they refer to others in our social environment—and are thus good candidates to explain judgment and behavior in moral situations.

Note that simple heuristics are applicable to moral as well as to non-moral problems. They thus contrast with the idea that people follow specific moral rules as characterized by moral philosophy or rely on "moral heuristics" (Sunstein, 2005) or an innate "moral grammar" (Hauser, 2006; Mikhail, 2007). Instead, moral judgment and behavior may often be a function of more general simple heuristics, a position in line with social and evolutionary psychology (Haidt & Kesebir, 2010; Krebs, 2008).

The goal of our investigation is neither to provide a normative moral theory nor to add to research on moral reasoning and behavior. Instead, we outline a descriptive approach that may prove useful for understanding the causes of judgment and behavior in

moral situations even when people do not rely on explicit moral reasons. Thus, we approach the question of why people fail to apply moral values from the more general perspective of how people make judgments and decisions in the social domain, of which moral situations are only a subset. Given certain ethical[4] values a society wants to uphold, this approach allows for recommendations on how to design environments that actually promote these values. To outline the main idea behind our first hypothesis, we selectively contrast it with three theories that attribute moral judgment and behavior to internal causes such as character traits, moral rules, or social preferences. We ask how these theories can account for three types of moral inconsistencies that people show, and argue that the interaction of heuristics and environment may provide a more parsimonious explanation. In the second part of the chapter, we elaborate the thesis that the heuristics determining much of judgment and behavior in moral situations are not specifically *moral* rules, but morally neutral heuristics that serve certain social goals. We end by outlining constraints and implications that follow from this perspective.

Beyond Character: How to Explain Three Moral Inconsistencies?

The road to hell is paved with good intentions.

Proverb

Understanding the causes of judgment and behavior in moral situations is ever more important in light of the myriad moral abominations that this young millennium has already seen; for instance, the torture and prisoner abuse committed by U.S. soldiers at Abu Ghraib, the widespread cases of child abuse by Catholic priests, or recent large-scale scams in the world of finance. In many of these cases, it is quite natural for us to locate the causes within the perpetrator by attributing moral or immoral behavior to a certain personality or character trait or lack thereof. Being sued for fraud for having intentionally bet against their customers, the leaders of Goldman Sachs may be seen as dishonest, greedy, or selfish, and these character traits are in turn used to explain their behavior. Assumptions about moral or immoral character ubiquitously emerge in public discussions and can have potentially serious consequences; for instance, by implying that only detention can protect the general public from habitual offenders. Not surprisingly, explaining moral behavior in

[4] We will use the terms ethical and moral interchangeably.

terms of internal traits is also prominent in influential theories across different disciplines. From virtue theory, to Kohlberg's developmental theory of moral reasoning (Kohlberg, 1984), to current theories of inequity aversion in economics: who you are supposedly determines what you will do.

However, focusing on internal traits can obstruct our view of the external world. Consider three kinds of moral inconsistencies commonly displayed that are not easily explained in terms of traits: inconsistencies in moral behavior across situations, inconsistencies between moral judgment and reasoning, and inconsistencies between moral judgment and behavior. Behavioral inconsistencies across situations occur when seemingly ordinary people do bad—sometimes even horrific—things. Although they may have behaved ethically in the past, people may suddenly and unexpectedly fail to do so in other similar circumstances. In 1994, more than 800,000 Rwandans were murdered when the Hutu systematically started to slaughter the Tutsi minority population with whom many of them had previously lived and worked: "Neighbors hacked neighbors to death in their homes, and colleagues hacked colleagues to death in their workplaces. Priests killed their parishioners, and elementary school teachers killed their students" (Gourevitch, 1995). Inconsistencies between moral judgment and reasoning occur when people falsely believe their judgments to be based on particular reasons; being invalid, these reasons provide merely a post hoc justification (Haidt, 2001). Finally, judgment and behavior are inconsistent when people's actions do not match their own proclaimed values and judgments. For instance, many people would endorse the statement that one should not lie, yet people lie on average three times in the first ten minutes when becoming acquainted (Feldman, 2009).

Any explanations of moral judgment and behavior needs to be able to account for such inconsistencies—yet these inconsistencies pose a problem from a trait perspective. If internal stable qualities explained judgments and behavior, how could one account for morally abominable behavior displayed by people who are otherwise quite ordinary? As Hannah Arendt (1964, p. 276) remarked, the "trouble with Eichmann was precisely that so many were like him, and that the many were neither perverted nor sadistic, that they were, and still are, terribly and terrifyingly normal." From a trait perspective of behavior, the observed inconsistencies raise the question of what prevented people from acting in a morally consistent way across situations.

In order to answer these questions, one strand of research in moral judgment and decision making has focused on biases that impede an ethical decision-making process (Messick & Bazerman, 2001) and on bounded ethicality (Chugh, Bazerman, & Banaji, 2005). Another line has studied cases of ethical blindness (Palazzo, Krings, & Hoffrage, 2012) and missing moral awareness (for a review, see Tenbrunsel & Smith-Crowe,2008), where people fail to see the moral dimension of a situation as a result of how they categorize it—and thus unintentionally fail to apply ethical standards in their judgments and decisions.

We add a further explanation by outlining how systematic inconsistencies in moral judgments and decisions are predicted by the interplay between a heuristic decision process and the environment. Extending the notion of ecological rationality (Gigerenzer et al., 1999), we propose the notion of *ecological morality* to point to a specific contribution of the environment to the resulting behavior: whereas internal dispositions or traits predict one kind of behavior for an entire range of situations, the same heuristic predicts different outcomes, ethical or unethical, depending on differences in the environment. We suggest that this dependency on the environment may account for evidence that poses a problem for internal explanations of moral behavior, without reference to biases or a state of missing moral awareness.

Inconsistencies in Moral Behavior Across Situations

Let us start with behavioral inconsistencies that people show across situations. How can a theory such as virtue ethics (Anscombe, 1958; Foot, 1978; MacIntyre, 1981) account for them? Virtue ethical accounts are close to the character-based explanations that seem quite natural to many people. Prima facie, virtue ethics is a moral theory, making normative claims about what kind of person we should morally become, and thus is not designed to make predictions about human behavior. At the same time, it does make psychological assumptions. Unlike moral theories that focus on moral rules or consequences, virtue ethics gives priority to the notions of virtues and moral character in moral evaluation. According to the Aristotelian account, a virtuous character is determined by what we as human beings are and need—and virtuous character traits are those that promote *eudaimonia*; that is, human flourishing and well-being (Oakley, 1996). Someone with a virtuous character should have the corresponding attitudes “at the right times, with reference to the right objects,

towards the right people, with the right motive, and in the right way" (Aristotle, 1984). Consequently, virtues are often understood as robust traits or deep-seated dispositions that are stable over time and relevant situations—and thus involve empirical assumptions in the explanation of moral behavior (Doris, 2002; Harman, 1999). For instance, someone blessed with the virtue of patience should display patience with his or her superior at work but also when dealing with an intern or when waiting for the long-term payoff of an investment. The person does not need to exhibit the disposition in every situation but should at least demonstrate a stable behavioral pattern over time and across situations in order to be considered virtuous in this way.

It has been one of the major lessons of social psychology and situationism that such cross-situational behavioral consistency is less common than one may expect. Personality traits have not been found to be very predictive for behavior across situations (Mischel, 1968; Ross & Nisbett, 1991), and to the extent that virtues can be seen as personality traits, the same critique applies to them (Harman, 1999; Doris, 2002; but see Prinz, 2009). A robust pattern of findings in social psychology demonstrates how easily personality traits are overwritten by even small changes in the environment. In a classical field experiment with theology students who train for a life guided by charity and thus should be expected to help someone in need, Darley and Batson (1973) found even these students not to be "good Samaritans," easily distracted from helping a person in need by a small situational manipulation. In a high-hurry condition, when the experiment required them to change buildings partway through, only 10% of the students offered help to a (confederate) person slumped in a doorway, compared to 63% in a condition of low time pressure. The change in behavior appears quite disproportionate to its—morally irrelevant—cause.

Situational influence also extends beyond "sins" of omissions to cases of active harm. Two of the most famous studies are Zimbardo's Stanford Prison Experiment (Zimbardo, 2008) and Milgram's obedience studies (Milgram, 1963; 1974). In one of Milgram's experiments, the experimenter instructed the participant to administer (as they thought) electric shocks of increasing intensity whenever a person gave an incorrect answer in a learning task. No fewer than 83% of the participants went beyond the 150-volt level, and 65% even continued to give shocks up the level of 450 volts in 15-volt steps. Although Elms and Milgram (1966) reported that obedient participants scored significantly higher on a scale for authoritarianism, they did not find any standard personality variable

that varied between "defiant" and "obedient" subject; in his review, Blass (1991) concluded that, although there was a lack of predictable patterns in situational manipulations, results on theoretically based links between personality factors and obedience were also weak and mixed. Apart from the unsettling result that even small situational changes led to unethical behavior in the absence of coercion, it is the ordinariness of the people, randomly assigned and tested in these experiments and the numerous replications, that speaks against an explanation by traits and for the power of the situation. More than 40 years after Milgram's classic investigations, this power can still be observed: in a partial replication up to the 150-volt level in 2009, 70% of the participants obeyed the instructions despite having been explicitly told that they could withdraw without losing their participation fee (Burger, 2009). As Milgram concluded, "often it is not so much the kind of person a man is as the kind of situation in which he finds himself that determines how he will act" (Milgram, 1974, p.205).

Proponents of virtues or other personality accounts could consider various ways to explain such behavioral inconsistencies. They could posit different traits for different situations—yet any ad hoc or inflationary account of such traits seriously undermines their explanatory value. They could rely on some aggregate conception of personality traits—and give up on the idea of predicting concrete behaviors (Epstein, 1979). Or they could try to identify overlooked cognitive or situational moderators that prevent people from judging and acting according to their moral personality, as is done by person-situation-interactionist accounts (Trevino, 1986). Although person-situation-interactionist accounts show that situations matter, analyzing the variance explained by personality and situational factors does not tell us much about what situations matter, how and why. However, specifying the heuristic processes will allow for clear predictions of what environmental differences matter and thus facilitate our understanding of the dependency of behavior on the situation.

How the interplay of process and environment explains apparent inconsistencies. From the perspective of ecological morality, cross-situational inconsistencies are to be expected. First, the same heuristic can lead to different outcomes, depending on differences between environments. Second, different heuristics may be selected based on properties of the environment.

Let us start with the first case, where inconsistencies are explained by the interaction of process and environment. The study on parental investment (Hertwig et al., 2002) illustrates how the same heuristic can lead to different outcomes, based on a difference in the environment. Organ donation is another case in point: between 1995 and 2003, on average 5,000 Americans and 1,000 Germans a year died while waiting for a suitable organ to be donated. Although most citizens say that they approve of organ donation, relatively few sign up as a donor: until 2002, only about 28% and 12% in the United States and Germany, respectively. In contrast, 99.9% of the French and Austrians are potential donors. Explanations along the lines of traits such as selfishness, fear, or empathy cannot account for the huge differences in donor rates unless one assumes enormous intercultural variation in these traits. That citizens of some countries are less informed about the necessity for organ donation also provides only a weak explanation: an exhaustive campaign in the Netherlands with 12 million letters sent out to a population of 16 million did not make a large difference. In a survey, 70% of the Dutch said they would like to receive an organ from someone who has died, should it be necessary, and merely 16% said they were unwilling to donate (Persijn & Van Netten, 1997). Nevertheless, only 27.5% signed up as a donor. The enormous variations among countries in donor rates across can, however, be explained by assuming that most people rely on the same heuristic, the *default heuristic*:

If there is a default, do nothing about it.

As Figure 2 shows, this heuristic will lead to different outcomes when environments differ in their legal defaults (Johnson & Goldstein, 2003).

In the United States[5], Germany, and the Netherlands, by law nobody is a donor unless one opts in. In France and Austria, the default is the opposite: everyone is a donor unless one opts out. Thus, the difference in potential organ donors depends on whether the default presumes consent or not.[6]

The *equity* and *default heuristics* illustrate how inconsistencies in behavior across moral situations can be understood and predicted. The same is true for *tit-for-tat.*

[5] In some countries, the policy is not the same nationwide. In the United States, for instance, some states have an opt-in policy, whereas others require citizens to make a choice.

[6] Note that presumed consent alone cannot fully explain the variation in donation rates across countries (Rithalia, McDaid, Suekarran, Myers, & Sowden, 2009). However, a cross-country study found that presumed consent legislation has a sizable effect after controlling for other determinates (Pettit & Knobe, 2009).

This simple heuristic, applicable to strategic interactions between two people, can also lead to apparently inconsistent behavior:

> *Cooperate on the first move of an interaction. Afterwards, always imitate the last action of the other.*

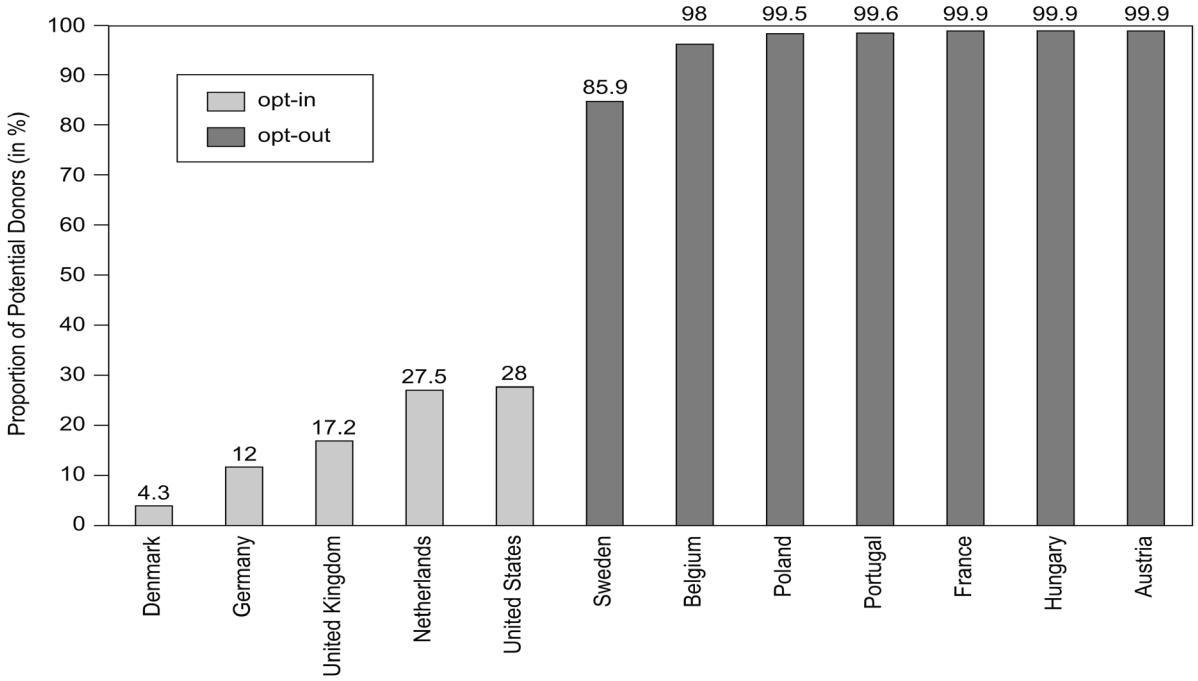


Figure 2. Why are so few citizens in Denmark, Germany, the United Kingdom, the Netherlands, and the United States potential organ donors? The percentage of potential organ donors varies between countries with opt-in and opt-out policies. In the United States, some states have an opt-in policy, whereas others require citizens to make a choice. (Source: Adapted from figure "Effective consent rates, by country" in "Do defaults save lives?" by E. J. Johnson & D. Goldstein, Science, 302, 1338–1339. Copyright 2003 by Science.)

Imagine a district that hires a new female prosecutor. Whereas the woman turns out to be cooperative and supportive while dealing with one advocate, she is never cooperative in cases defended by a second one. Dealing with the first advocate, she appears to be a nice person, and in interactions with the other advocate, a nasty one. However, this seeming capriciousness does not necessarily mean that the female prosecutor is inconsistent. She may rely on the same heuristic in both cases, tit-for-tat, and her varying behavior may simply reflect a difference in her two colleagues' behavior. While the first advocate had been cooperative and supportive to the prosecutor when she was new, the second was at first skeptical and reluctant to accept her; later on, he found her behavior to prove him

right. By relying only on the last action of another person, tit-for-tat ignores most of the history of interactions with other people. And depending on the strategies other people use, tit-for-tat can lead to a continuing circle of noncooperation. Nevertheless, under certain conditions, it can lead to more cooperation and better outcomes than the rational strategy "always-defect" (e.g., in a Prisoner's Dilemma with a fixed or undetermined number of trials).

However, behavior in moral situations may also depend on the environment in a second way, given that different heuristics may be selected based on properties of the environment. Imagine a person working for a company who gets the chance to be the team leader of a large project and has to decide how to distribute responsibilities among the members of the team. Assume that this person has only three heuristics at his disposal: *1/N,* the *default heuristic*, and *imitate-the-successful*; the latter is a heuristic that adopts the behavior of a successful person (here, another team leader). Because only two other colleagues were assigned such a large team before, no default exists concerning how responsibilities should be shared. Among the three heuristics, only two are thus applicable to the problem at hand. In line with flat hierarchies common in the company, the team leader may share the workload and responsibilities equally among all team members. Yet he may also decide to imitate one of the previous team leaders, who was quite successful by implementing a strict hierarchical order in which he controlled all steps personally. Which heuristic will be selected might be determined in an adaptive way, based on features of the environment (Payne, Bettman, & Johnson, 1993). If the relevant differences in the environment are small or go unnoticed, different heuristics may be selected and the behaviors appear inconsistent across situations. This process of selection could, for instance, be conceptualized in terms of reinforcement learning within strategy-selection theory (Rieskamp & Otto, 2006), and individual differences may exist in the heuristics selected. However, for strategies that people use in making inferences, Bröder (2012) tested a variety of personality measures and found no stable personality trait to be substantially correlated with the applied strategies. Instead, when testing for individual differences in cognitive capacities, he found higher intelligence scores to be related to a more adaptive strategy selection. Although social goals will, of course, be different from those in the cognitive domain, we suggest that social intelligence may analogously be understood as strategy selection adapted to the multiple goals that humans pursue within the social domain.

To summarize: If one assumes personality traits such as virtues or empathy, cross-situational inconsistencies in behavior are difficult to explain. Person-situation-interactionist accounts of missing moral awareness Butterfield, Trevino, & Weaver, 2000) or ethical blindness (Palazzo et al., 2012) offer an explanation for those inconsistencies that can be regarded as cases of "unintended unethicality" (Tenbrunsel & Smith-Crowe, 2008). They do so by taking into consideration properties of the environment that may make people fail to see the moral impact of a situation—and, as a consequence, fail to apply their own moral values. In contrast to such accounts, the concept of ecological morality requires specifying the decision processes in order to understand the interaction with different environments and to predict situational variation accordingly. This allows for a different explanation of unintended unethical as well as ethical behavior. The equity heuristic, for instance, which parents may consider to be a just distribution strategy, can unintentionally lead to a result judged to be unfair or immoral, depending on the number of children. And a decision-making process like the default heuristic, which makes no reference to ethical considerations, can nonetheless lead to both immoral and moral outcomes. The interaction of process and environment thus provides a parsimonious explanation for behavior that appears inconsistent from an internal point of view—without reference to moral awareness and like concepts. Such an explanation, however, needs to be supplemented by a theory of how strategies are selected according to properties of the environment.

Why inequity aversion cannot explain inconsistencies in moral behavior across situations. A class of theories that has become popular in behavioral economics explains certain kinds of moral behavior by taking into account social preferences such as altruism, fairness, or reciprocity. These theories differ from economic models that assume that people maximize their utility and are solely motivated by material self-interest. One prominent example used to explain apparently unselfish behavior in social interactions is the theory of inequity aversion Fehr & Schmidt, 1999), which integrates an individual's aversion to unequal outcomes into the utility function. In contrast to virtue ethics or approaches inspired by Kohlberg (1984), the theory thus combines the idea of personality features with the calculation of the expected utility of an action:

$$U_i(x) = x_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max\{x_j - x_i, 0\} - \beta_i \frac{1}{n-1} \sum_{j \neq i} \max\{x_i - x_j, 0\},$$

where the utility of an option for an individual *i* is characterized by two parameters α_i and β_i that are supposed to capture *i*'s degree of inequity aversion. While α_i measures *i*'s aversion to being poorer than others, β_i refers to *i*'s discomfort at being richer. An individual's equality preferences α and β are estimated from his or her choices in economic games such as the ultimatum game. These games represent different social situations in which the outcome of a person's decisions depends on the decisions of others. The goal is to find a utility function that accounts for behavior across more than one game (Camerer, 2003).

Given common behavioral inconsistencies across situations, however, what are the options for inequity aversion and similar theories?[7] Inequity aversion makes two critical assumptions that pose a problem for the explanation of behavioral inconsistencies. First, it resembles *prospect theory* (Kahneman & Tversky, 1979) or *utilitarianism* in that it relies on optimization: the individual is predicted to choose the option that maximizes some form of utility. To determine the optimal choice, these theories require "small worlds" (Binmore, 2008; Savage, 1954), where all outcomes, consequences, and their probabilities are known and no surprises happen. In the real world, such knowledge is missing most of the time, and optimization is unfeasible. Consequently, the inequity aversion model is assumed to be no more than an *as-if* model (Friedman, 1953; Berg & Gigerenzer, 2010), which is mute on the actual cognitive processes that individuals follow. Second, inequity aversion assumes that individuals have concrete preferences about equality issues. In short, the theory holds that these individuals decide as if they have maximized their utility, given individual equality preferences. Although the theory is formulated in terms of individuals who differ in their equality parameters, Fehr and Schmidt (1999) tested the theory solely in terms of aggregate distribution of choices. This led to a heated controversy (Binmore & Shaked, 2010; Fehr & Schmidt, 2010). Yet the theory not only assumes the aggregate dis-

[7] Behavioral economists do not assume that behavior observed in economic games gives rise to externally valid predictions for real-world situations. For instance, behavior in dictator games is not assumed to predict how much people will donate to someone on the street (Binmore & Shaked, 2010; Levitt & List, 2007). However, this restriction to internal validity is not always made clear (e.g. Ariely, 2008; Thaler & Sunstein, 2008). Yet, even across economic games, the evidence for stable preferences seems mixed (Andreoni & Miller, 2002; Binmore & Shaked, 2010; Blanco, Engelmann, & Normann, 2010; Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999).

tribution of parameters to be stable, but also predicts cross-situational correlations of individual behavior based on the estimated utility function (Fehr & Schmidt, 1999, p. 847)—and thus assumes some kind of stability of social preferences across games, as does trait psychology. However, in within-subjects tests across different games, inequity aversion had only low predictive power (Blanco, Engelmann, & Normann, 2010).

These assumptions have consequences for the explanation of behavioral inconsistencies across situations. As an illustration, consider how *inequity aversion* could account for the behavioral variation in parental investment (Figure 1). Parents face a basic distribution problem of how to allocate care time to two, three, or four children at a certain point in time. Assume that a parent's equality preferences also cover non–self-centered equality or inequality among third parties—in this case, their children. The distribution pattern shown in Figure 1 differs systematically depending on the number of children, birth rank, and child-spacing. Thus, the distribution of equality preferences for parents with two children will not be consistent with the distribution for parents with three or four children. But is it plausible to assume that parents' equality parameters differ for two, three, or four children, or across the time span in which they are raising them? Or is it likely that the utility of middle-born children is smaller for parents than that of first- or late-born children?

To account for such patterns that are not consistent across situations, the theory of inequity aversion can be adjusted only by adding further parameters to the model, by referring to properties of the environment to explain the differences, or by varying the parameter values for equality preferences in order to identify a better distribution of parameters that fits across situations. Yet if the parameter values do not enable predictions for similar situations, the explanation amounts to no more than ad hoc data-fitting (Binmore & Shaked, 2010).

Alternatively, the theory may use properties of the environment to account for the differences across situations. Because inequity aversion is an as-if theory, however, it does not allow researchers to analyze the interaction of cognitive processes with environmental properties. Whereas the inequity aversion model can include environmental properties solely as further free parameters of the rational calculation, the equity heuristic does specify a process that takes the number of children as input and accordingly predicts variations in behavior without any free parameters. Systematic and predictable inconsistencies across situations thus follow directly from the heuristic.

To summarize: Even small changes in the environment have been found to influence behavior, making cross-situational consistency of behavior less common than one may expect. Virtuous dispositions, moral personality features, or other-regarding preferences predict only one kind of behavior for a set of situations and thus run into problems when facing systematic inconsistencies in moral behavior. The concept of ecological morality focuses on the interaction of heuristics and the environment instead: The same heuristic may lead to different outcomes, ethical or unethical, depending on the environment. This implies an explanation for behavior that appears morally inconsistent from the perspective of accounts that rely on internal traits and ignore the interaction of cognitive processes and environment. Specifying the simple non-optimizing processes facilitates understanding which environments they are adapted to and thus understanding why and when humans may succeed or fail in pursuing ethical goals.

Inconsistencies Between Moral Judgment and Reasoning

A second inconsistency that may be explained from the perspective of ecological morality concerns the (mis)match between the reasons we give for our moral judgments and the actual processes underlying them. To illustrate the issue, let us consider Kohlberg's (1984) influential theory of personality development. Kohlberg focused on "moral judgments," understood as the deliberate, rational application of moral criteria. The ability to apply these criteria is assumed to develop with age, according to Kohlberg's six moral stages: depending on the moral stage people have reached, their judgments are based on the criteria available to them, ranging from authority-based reasons to a post-conventional, impartial Kantian principle.

Recent research on moral intuitions has challenged the deliberative core of Kohlberg's account—when judging certain moral dilemmas, people seem to rely on intuition rather than on deliberative rational reasoning (Haidt, 2001). A mismatch often exists between the processes underlying moral judgments and their justifications. The reasons we give may not be the causes of our moral intuitions but rather a post hoc rationalization—without our being aware of it (Hauser, Cushman, Young, Kang-Xing Jin, & Mikhail, 2007; Nisbett & Wilson, 1977). In line with this, we do not believe that rational conscious reasoning is the paradigmatic process underlying moral judgments; instead, moral judgments may often be produced by heuristics.

How the judgment process and people's description of it can diverge is shown in a study by Dhami (2003), who investigated bail decisions made in two London courts over a four-month period. Bail decisions do not concern the defendants' guilt, but their trustworthiness. In the English system, 99.7% of the magistrates responsible for these decisions are community members without legal training. They decide whether to release a defendant on bail or to make a punitive decision, such as custody or imprisonment. In evaluating trustworthiness, they are supposed to consider several pieces of information, such as the nature and severity of the offense the defendant is charged with, personal characteristics, community ties, previous bail record, and so forth. When asked how they make their decisions, magistrates typically answered that they examine the full evidence carefully in order to treat the defendant fairly and to take into account the details of the individual case (Dhami & Ayton, 2001). This answer echoes the requirements set by the Bail Act of 1976. The story, however, looks different once the judgment process is modeled according to the actual decisions that magistrates made (Dhami, 2003). For Court A, 92% of 159 bail decisions could be predicted correctly by a simple fast-and-frugal tree: it models the magistrates' decisions as relying sequentially on requests made by the prosecution, the decisions made by the previous court, and the police (Figure 3). If any of these three opted for a punitive decision, the magistrates did as well.[8] As a consequence, their judgments varied depending on relevant differences in the environment; that is, as a function of the cue values that entered their judgments.

The judgment process modeled by the fast-and-frugal tree is inconsistent with how magistrates believed they made their judgments. But what explains this inconsistency between the processes underlying their judgments and the descriptions they give?

From the perspective of the Kohlberg model, judgments are a function of the developmental stage of the individual, which determines the available moral criteria. Consequently, people at the highest moral stages should be expected to base their judgments on reasons they are able to refer to in their justifications, independent of the judgment situation. Yet although the magistrates knew what they were supposed to do and even believed they had fulfilled their responsibilities, the successful model of their decisions in terms of a

[8] For Court B, the predictive accuracy was 85%, based on a tree that sequentially checks whether previous courts made punitive decisions, whether the police did, and, finally, whether the defendant was previously convicted for a similar offense. In comparison, Franklin's rule, a model that takes all information into account, predicted only 86% of the decisions for Court A and 73% of the decisions for Court B. Note that these values refer to prediction (cross-validation) and not simply to fitting data.

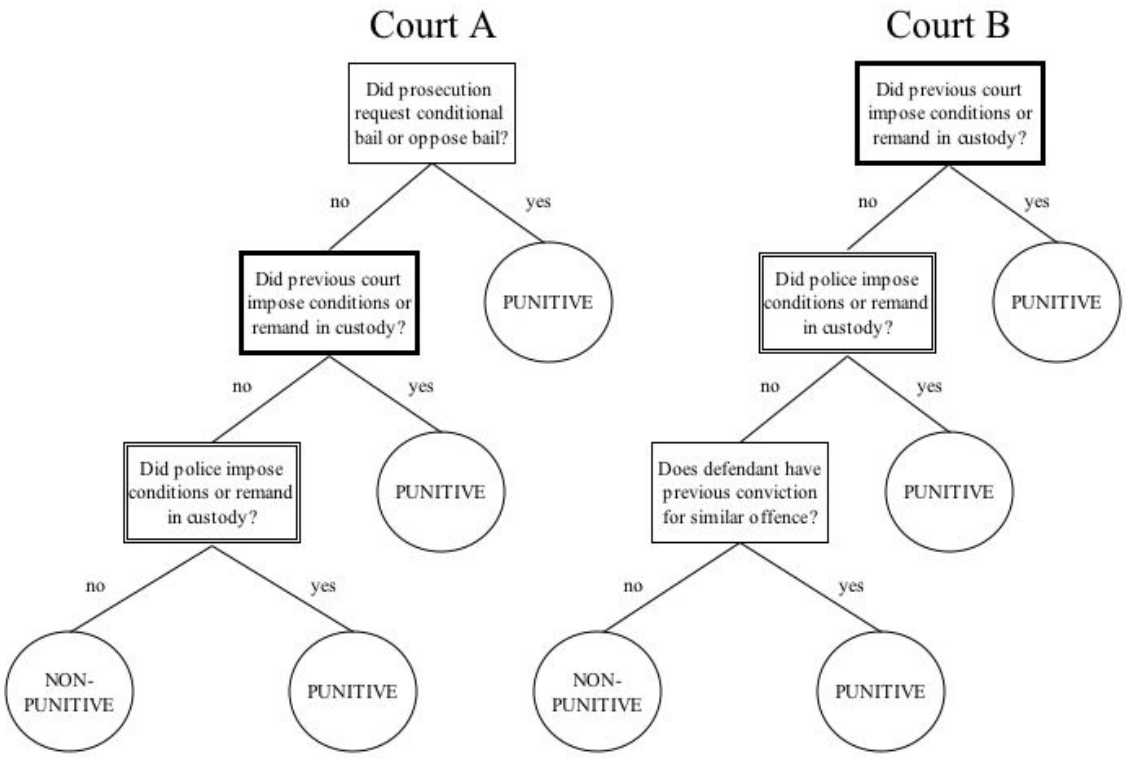


Figure 3. Models of fast-and-frugal heuristics for bail decisions in two London courts. (Source: Adapted from Figure 1 in "Psychological models of professional decision making" by M. K. Dhami, Psychological Science, 14, 175–180. Copyright 2003 by the American Psychological Society.)

simple heuristic (Figure 3) suggests a different judgment process. From an ecological perspective, it is important to go beyond the internal perspective because the heuristic adopted by magistrates may not be independent of the properties of the judgment situation. First, to consider and integrate all available information may well be an impossible task, given that the duration of bail hearings was, on average, only 6.7 minutes for each defendant. Under time pressure, people tend to use fewer cues and to rely on noncompensatory strategies (Rieskamp & Hoffrage, 1999, 2008). Second, the situation permits only one kind of error to be detected: magistrates do not receive information about whether they falsely jailed defendants, but only whether they falsely bailed someone who subsequently committed a crime. Given this asymmetrical feedback, magistrates are likely to adopt a defensive decision procedure of "passing the buck" (Dhami, 2003), even if it is inconsistent with how they believe they make their judgments and even if it violates the ideal of due process, which entails considering all relevant information about the defendant. Thus, features of

the environment, such as uncertainty due to asymmetrical feedback, appear to select the judgment process that magistrates apply, while their portrayal of the judgment process also serves the social function of justification (Lerner & Tetlock, 1999). Therefore, it may not be surprising that it does not match the actual processes.

In sum, the reasons that people give for a judgment may often not veridically reflect its true causes. From the point of view of ecological morality, inconsistency between the heuristics underlying judgments and the purported reasons for these judgments is to be expected. This does not mean that judgments never rely on reasoning processes and that deliberative moral reasoning does not have its place (Blasi, 2009; Haidt & Bjorklund, 2008). Yet we do not consider it helpful to explain the inconsistencies by a dichotomy of moral intuition versus reasoning, as done by dual-process theories (see Gigerenzer & Regier, 1996; Kruglanski & Gigerenzer, 2011). Once the processes are specified, the dichotomy of intuition and reason does not provide any additional information. Moreover, the division is too simplistic: intuitive and deliberative judgments can be based on the same processes, and conscious processes do not need to be rational or moral.

Inconsistencies Between Moral Judgment and Behavior

The third and last inconsistency that we discuss concerns the relationship between moral judgment and behavior, or moral cognition and action: people often make decisions that do not match their own moral judgments and values—without being aware of it (Chugh et al., 2005; Chugh & Bazerman, 2007; Palazzo et al., 2010). The gap between moral cognition and behavior has been challenging for Kohlberg's theory and its proponents, in which moral judgments are seen as a necessary precondition of moral action; as a consequence, behavior that does not rely on a moral judgment will not count as moral (Kohlberg, Levine, & Hewer, 1983). Hence, the stage of individuals' moral development should, at least in general, guide not only their moral judgment but also their moral behavior. The same assumption can be found in business ethics: in the common four-stage-model, moral awareness is assumed to precede moral judgment, followed by moral intention and finally moral behavior (Jones, 1991; Rest, 1986; Trevino, 1986). Consequently, much of this research focused on moral awareness as an important precondition for moral judgment *and* behavior.

However, moral awareness or developmental stages of moral judgments do not necessarily translate into actions. Admittedly, Kohlberg found that the tendency to help a drugged student increased parallel to the stage of moral judgment (Kohlberg & Candee, 1978), and that people displaying disobedience in a Milgram study had reached higher moral stages (Kohlberg, 1984) than those who obeyed. Yet even researchers supportive of Kohlberg's approach concluded that there is not much support for the thesis that people who have reached the post-conventional stage are more likely to resist social pressures in their moral actions than people at lower stages (Blasi, 1980). Moral stages are at best loosely related to specific behaviors. Criteria from different stages can lead to the same behavior, and the same stage can give rise to different behavioral predictions (Blasi, 1980). Similarly, results in business ethics concerning the relationship of moral awareness and moral behavior have been mixed (Tenbrunsel & Smith-Crowe, 2008). In order to explain the inconsistency between moral judgment and behavior, these internal approaches can only recruit further personal characteristics unrelated to moral cognition, concepts such as moral identity and situational moderators—or try to spell out the process that may possibly connect moral judgment and action.

In contrast, the perspective of an ecological morality does not presuppose consistency between moral judgments and behavior. The reason is that the situations in which we make a judgment are likely to differ from those in which we face moral choices. Consider the case of American teenagers who publicly take a vow of abstinence, pledging not to have sex before marriage. Typically coming from religious backgrounds, these teenagers uphold virginity as a moral value. As a consequence, one should expect these values to guide their behavior, even more so after they publicly declare them. Yet these teenagers were just as likely to engage in premarital sex (Rosenbaum, 2009) and, unlike those who did not take the vow, they were less likely to use condoms or other forms of contraception. However, if we think of teenagers' behavior as being guided by a heuristic such as *imitate-your-peers*, even a public pledge taken in one situation will not rule their behavior when it comes to sexual relations. Furthermore, if the heuristic is unconscious but teenagers believe their behavior to be under conscious control, moral judgment and behavior can diverge. Although teenagers may rely on the same simple heuristic in a situation where they judge virginity to be a moral value and in a situation where they face the choice of engaging in premarital sex, both situations differ with regard to what most of their peers do.

Thus, imitation of one's peer group can easily swamp moral statements that were made before.

Again, we do not want to suggest that moral judgment never guides moral behavior; yet nothing ensures that the processes underlying our judgments do in fact match those underlying behavior. Even if the process is the same, a change in the environment may result in different behavior. Thus, systematic inconsistencies between moral judgment and behavior can be expected. Moreover, nothing speaks against the idea that different processes and criteria become available at different points of development and socialization (Krebs & Denton, 2005). Yet, because moral theories were not meant to be predictive of behavior, starting from normative ethical assumptions will rule out an important range of processes underlying people's judgments and choices in moral situations (Lapsley & Hill, 2008). In contrast, the notion of ecological morality does not presuppose ethical standards that restrict the domain of enquiry, but conceptualizes behavior in moral situations from the functional perspective of what it serves and how it could have been evolved.

Which Heuristics Guide Judgment and Behavior in Moral Situations?

It has been suggested that the rules underlying people's judgment and behavior in moral situations are, in some sense, particular to the moral domain. Different views exist on what may constitute such a rule, none of which we believe to be a very promising explanation for judgment and behavior in moral situations. The first view holds that people rely on moral rules, such as the Ten Commandments. Moral rules are thus deontological rules for judgment and behavior, such as "Do not kill," or those that proscribe torture or the willful destruction of property. In a second view, people follow "moral" heuristics, such as "Do not tamper with nature" or "Punish and do not reward betrayals of trust" (Sunstein, 2005). According to this account, moral heuristics are understood as a wider set of moral intuitions underlying judgment and behavior, which do not necessarily match those specified by any normative moral theory. Finally, the view has been proposed that people follow different instantiations of an innate moral grammar, operating on content-free principles related to general characteristics of actions, such as consequences, causality, intentions, and responsibility, on which intuitive moral judgments across cultures are based (e.g., Hauser, 2006; Mikhail, 2007).

Why Moral Rules Are Not Good Explanatory Candidates

We argue that rules or heuristics particular to the moral domain often provide only weak explanations for people's judgment and behavior, for three reasons. First, moral rules, such as "Do not kill" or "Do not harm others" appear to describe reality insufficiently. More than 100 million people died at the hands of others during the twentieth century (Katz, 1993); and, in 2009, more than 43 million were refugees escaping from violence and war—more than ever in the last 15 years (UNHCR, 2010). Both numbers are powerful reminders that people routinely do not act according to general moral rules that ban violence. Attempting to describe behavior by such moral rules would require telling a plausible story that explains not only the particular exemptions that we make, for instance, in times of war or in the legal context (e.g., death penalties) but also why moral rules can so suddenly be disabled, as in the civil wars in Rwanda or the former Yugoslavia.

Second, the question of what counts as a moral rule cannot be separated from the delineation of the moral domain; that is, from which questions are actually considered to be "moral" at all. Starting from an empirical definition, what counts as a moral issue changes not only across cultures but also throughout history. Whereas contemporary Western views focus on issues of individual rights and harm, other values are seen as more important at different times and places (Haidt & Joseph, 2007). When the United States established equal rights for everyone, "everyone" did not include women or a large proportion of slaves who were not given the same rights. Even within the same country, norms intended to secure individual rights often coexist with conflicting "honor killings" related to family, authority rights, or religious norms. In 2002, the religious police in Saudi Arabia prevented male rescue workers from saving 15 female students trying to escape from a burning school who were not wearing the headscarves and black robes required by the religious norms of the country (Human Rights Watch, 2002). Given these differences in moral systems, an empirical investigation of moral judgments and behavior should be careful not to restrict itself *a priori* to a Western-biased delineation of what counts as moral (Haidt, 2007; Shweder, Much, Mahaptra, & Park, 1997).

There is a third reason to consider heuristics to be morally neutral: the same heuristics are used on both sides of the "moral rim." Issues such as smoking (Rozin, 1999) or the use of renewable energy are examples of behaviors that only recently have become moralized. The heuristic in operation, however, may have remained the same. The default heu-

ristic is a case in point: a natural experiment in the German town Schönau showed that when "green electricity"[9] was introduced as the default, almost all citizens went with the default, even though nearly half of them had opposed its introduction. In contrast, in towns with "gray" energy as the default, only about 1% chose green energy. This pattern was replicated in laboratory experiments (Pichert & Katsikopoulos, 2008). The default heuristic also appears to guide behavior in domains that are typically not considered moral situations according to Western standards, such as the purchase of an insurance policy. In Pennsylvania and New Jersey, drivers are offered the choice between an insurance policy with unrestricted right to sue and a cheaper one with restrictions on the right to sue (Johnson, Hershey, Meszaros, & Kunreuther, 1993). The unrestricted policy is the default in Pennsylvania, and the restricted one the default in New Jersey. If drivers rely on the respective defaults, more drivers in Pennsylvania should be expected to buy the unrestricted and more expensive policy. And, in fact, 79% of the Pennsylvanian drivers bought the unrestricted policy, whereas only 30% of the New Jersey drivers did so.

Thus, from organ donation and environmentally friendly behavior to non-moral activities such as the choice between different insurance policies, there is evidence that people rely on defaults in their choices—and thus that the same heuristic is used inside and outside the moral domain. Given the inconsistencies between moral rules and people's behavior, the cultural and historical variations in the delineation of the moral domain, and the fact that the same heuristic can be used on both sides of the moral rim, we suggest that the heuristics that may often guide behavior in moral situations are typically morally neutral. The same heuristics may underlie moral and non-moral, social and nonsocial judgment and decision making, allowing for a parsimonious explanation across domains.

"Social" Heuristics as an Explanation of Judgment and Behavior in Moral Situations

If the heuristics underlying judgment and behavior are often morally neutral, which heuristics could then determine much of behavior in moral situations? From the adaptive-toolbox perspective, one way to approach this question is by looking at the goals that people pursue within a broader social and even evolutionary perspective (Haidt & Kesebir, 2010; Krebs, 2008): What was the original evolutionary function of morality in its different

[9] "Green electricity" is energy produced from natural sources, such as wind or solar power, that have fewer negative environmental impacts than burning coal or gas.

instantiations—and which heuristics are likely to fulfill this function? According to Darwin (1871), morality serves a social purpose that has evolutionary advantages—the social coherence of a group. We propose that heuristics serving this function guide our judgment and behavior in moral situations. For instance, the equity heuristic helps us avoid distribution conflicts; the default heuristic promotes consistent behavior within a group; and imitation processes strengthens the unity of groups, and can also ensure effective learning. Furthermore, there exist genuinely "social" heuristics that process social information (e.g., imitate-your-peers, tit-for-tat, or decision trees with social cues). Because they refer in some way to others whom we rely on, interact with, and depend on in our social environment, these heuristics can be expected to strongly further social coherence.

The importance of social relations for human behavior is supported by the fact that humans have been shown to have a distinctly social adaptation that allows them to enter a cultural world with its social practices and institutions. Humans are not the only ones who imitate, but they show uniquely sophisticated skills for social cognition, social learning, and communication (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007). People use information from their social environment when making decisions in moral situations. For instance, in repeated public-goods games in which people do not know how much others will contribute, experiencing a randomly assigned person who contributes before the others leads to higher contribution levels, even without any mechanisms to punish free-riders (Güth, Levati, Sutter, & Van der Heijden, 2007). Also, people appear to rely on the behavior of others in their interpretation of a situation. In a classic experiment showing how groups inhibit helping behavior, Latané and Rodin (1969) had undergraduates participate in what they were told was a market research study, conducted by a female researcher. While participants were filling out a questionnaire, the woman went behind a curtain to wait. In the company of an unresponsive confederate, only 7% attempted to help when they heard a loud crash and cries of pain, compared to 70% among those who were alone when the commotion occurred. In groups with two, not previously acquainted, persons, only in 40% of the groups did one person help at all. Although fewer than 5% of the participants thought that the cries were recorded (as they were), non-helpers claimed in post-experimental interviews that they would have helped if the emergency had been "real." Variations of Milgram's paradigm point in the same direction: many obedient participants showed strong signs of distress, indicating a conflict when following the requests of

the experimenter (Milgram, 1963), yet obedience rates dropped only when people sequentially witnessed two confederates' disobedient behavior (Milgram, 1974) and obedience virtually disappeared when two experimenters disagreed. Burger (2009), however, observed that one disobedient person was not enough to top reliance on authority. Thus, people do not follow their own evaluation of a moral situation; yet they do react to certain conflicting information provided by others.

By taking social information as input, heuristics such as imitate-your-peers, tit-for-tat, or the default heuristic (which is a form of institutionalized recommendation) may thus serve several competing social goals that people face at the same time: to solve the social task at hand while preserving a certain position within the group. Trying to explain condemnation mechanisms, DeScioli & Kurzban (2009) proposed understanding conscience, not as the correct internal application of moral norms, but rather as a defensive mechanism to avoid punishment or condemnation by others—conscience may be better understood as "the inner voice that warns us somebody may be looking" (p. 290).

However, we still lack a systematic theory of the structure of social environments that are relevant to social interactions and moral situations. For instance, how can we understand the many situational variations found to influence obedient behavior as a result of the interplay of heuristics and the environment? Why do people follow the request of the experimenter when witnessing one person being disobedient, but stop doing so when witnessing more disobedience? Research on social structures (Boyd & Richerson, 2009; Fiske, 1992; Haidt & Bjorklund, 2008; Shweder et al., 1997) could be one starting point. For instance, Fiske's (1992) four kinds of relationships—equality matching, market pricing, authority ranking, and community sharing relations—could constrain the set of heuristics that may be triggered. Authority ranking is characterized by a linear, asymmetrical ordering of persons. Cues indicating asymmetrical relations, such as spatiotemporal orders, may thus mainly select heuristics that use input from others who rank higher or lower.

Other important features that may influence the heuristics selected are the structure of social interconnections, which determines the position of a person within their social network, their interaction partners and information sources; and the stability of the social environment, resulting from the degree of institutionalization or the heterogeneity of the strategies used by others. Both the social structure and the stability of the social environment determine the degree of social uncertainty about what others will do and what is ap-

propriate—and may thus influence which heuristics actually foster social coherence and which ones people apply.

As an important methodological consequence, understanding moral judgment and behavior requires studying social settings that allow people to rely on the cues that may have a strong impact on their behavior—and that show the environmental and social uncertainty that is common in the real world. A decision may concern a moral issue, yet people may, in fact, apply simple heuristics that serve powerful social goals. An important aspect of the ecological morality perspective is that a heuristic may be socially rational, given these goals, while not judged to be ethically appropriate from a normative point of view.

Ecological Morality and Moral Theory

The perspective of ecological morality aims at understanding the heuristics that underlie people's behavior in moral situations. Prima facie, ecological morality and moral theory provide two separate views on the same behavior: one offers a descriptive explanation, the other a normative theory about what people should do according to an ethical standard. When the goal is to explain moral behavior, so we argue, it is not useful to constrain the domain of inquiry by a commitment to one normative moral theory or another. Instead, the starting point should be the set of issues that count as moral within a given society. Without committing to a specific moral theory, it is possible to investigate judgments and behavior in moral situations by focusing on the (in)consistencies we discussed: the consistency of behavior across situations, between the justifications people give for their intuitive judgments and the processes behind them, as well as between these judgment processes and moral behavior.

Although we believe it is useful to acknowledge different instantiations of what counts as moral, we do not want to promote ethical relativism. Nevertheless, the descriptive and the normative perspectives are not entirely disconnected from each other. There are three important points of contact:

(a) *Normative moral theories are based on psychology and culture*. The study of ecological morality does not presuppose any particular ethical goals but fits a variety of social goals that people pursue, as well as evolutionary functions that moralities may serve. If the heuristics that people follow are culturally and evolutionary adaptive, selection may, for instance, favor heuristics that promote group coordination—but selection is ethically

blind (Cosmides & Tooby, 2004). Nothing ensures that the resulting judgment and behavior are morally justifiable based on a moral theory that a society wants to adopt. Moreover, general "social" heuristics underlying judgments and behavior, such as imitation processes, may support the selection and transmission of shared moral intuitions and corresponding judgments (Simon, 1990). These intuitions vary significantly across groups and may develop into opposing moral norms, such as the "Do not kill" rule in twentieth- and twenty-first-century Western philosophy, and rules supporting killing in particular situations, for instance, sister-killing in Muslim families, heroism in warfare, or the death penalty in democracies such as the United States and India. Insofar as moral theories are built on pre-theoretical moral intuitions—as pointed out by experimental philosophy (Feltz & Cokely, 2009; Knobe & Nichols, 2007)—knowing about the causes that give rise to our shared moral intuitions and behavior should be seen as highly valuable for any moral theory. Empirical knowledge enables us to distance ourselves from the strength of our intuitions, to question their apparent universality, and to judge whether some intuitions produced by a simple heuristic may not fit the societal problems and values of today.

(b) *Normative moral theories need to satisfy psychological constraints*. Facts about the working of our minds are of crucial importance for a reasonable normative theory. First, psychological findings constrain assumptions and claims that normative theories presuppose; for instance, in the case of broad personality traits assumed by virtue ethics. Second, there should be no moral requirements that are in principle impossible for humans to fulfill. "Ought" implies "can": normative theories should neither call for decision procedures that humans will not be able to follow, nor should they suggest normative standards—as consequentialism risks doing—that can hardly be determined in the uncertainty of the real world (Gigerenzer & Sturm, 2012).

(c) *The implementation of ethical norms requires empirically plausible prescriptions.* Knowledge about psychological mechanisms and the environments that they work in is necessary for coming up with reasonable prescriptions that promote the ethical norms that a society considers to be essential. Questions of responsibility may arise if we acknowledge the impact of the environment on behavior. But it does not follow from this acknowledgement that we have to absolve unethical behavior or cannot hold people accountable for their actions. However, understanding the interplay of cognitive processes and environments opens up a chance of effectively promoting the ethical goals that a socie-

ty values—and it is an empirical question of how to design environments that effectively do so.

The outline that we have presented here is based on work that tries to understand how simple heuristics make us smart. In the context of moral situations, however, the question is: Do simple heuristics make us good? The answer is: No. Just as simple heuristics only perform well in some environments, the same holds true for heuristics in the moral domain. The study of bounded and ecological morality does not suggest that simple heuristics make us good. But knowing the heuristics in the adaptive toolbox, including what triggers their use, and designing the environments accordingly can make us behave better.

Chapter 3

Cooperation in Risky Environments: Decisions from Experience in a Stochastic Social Dilemma[1]

Florian Artinger [a], Nadine Fleischhut [b], Vittoria Levati [c], & Jeffrey R. Stevens [d]

[a] Warwick Business School, Warwick

[b] Max Planck Institute for Human Development, Berlin

[c] Faculty of Economics, University of Verona

[d] Department of Psychology, University of Nebraska-Lincoln

[1] Artinger, F., Fleischhut, N., Levati, V., & Stevens, J. (2012). Cooperation in risky environments: decisions from experience in a stochastic social dilemma. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 84-89). Austin, TX: Cognitive Science Society.

Abstract

Often in cooperative situations, many aspects of the decision-making environment are uncertain. We investigate how cooperation is shaped by the way information about risk is presented (from description or from experience) and by differences in risky environments. Drawing on research from risky choice, we compare choices in stochastic social dilemmas to those in lotteries with equivalent levels of risk. Cooperation rates in games vary with different levels of risk across decision situations with the same expected outcomes, thereby mimicking behavior in lotteries. Risk presentation, however, only affected choices in lotteries, not in stochastic games. Process data suggests that people respond less to probabilities in the stochastic social dilemmas than in the lotteries. The findings highlight how an uncertain environment shapes cooperation and call for models of the underlying decision processes.

Keywords: Decisions from Experience; Social Dilemma, Cooperation; Risky Choice; Public Good.

Introduction

When people face an opportunity to cooperate, such as when opening a business together or pursuing a joint research project, the outcomes of these enterprises are frequently uncertain. On the one hand, joint enterprises often constitute a social dilemma, where it is in the collective interest of the group to cooperate, yet individually rational to free ride. Despite these incentives, there is overwhelming evidence that many people still engage in cooperation (e.g., Ostrom, 1990; Ridley, 1996). On the other hand, even if people cooperate, outcomes often are uncertain due to a risky environment. For instance, even if all business partners cooperate, a new start-up may fail due to external events, such as natural disasters disrupting supplier shipments. Laboratory experiments show that when social dilemmas are embedded in a stochastic environment, cooperation declines sharply (for a review see E. Van Dijk, Wit, Wilke, & Budescu, 2004). What has not been addressed is how different levels of environmental risk and the format in which it is presented affect cooperation.

Studies on risky choice find a pronounced difference in behavior depending on how information in lotteries is presented: whether people sample the distribution of outcomes (*decisions from experience*) or decide based on a summary description of outcomes and probabilities (*decision from description*) (Hertwig, Barron, Weber, & Erev, 2004; for a review see Hertwig & Erev, 2009; Rakow & Newell, 2010). In conventional lotteries with described probabilities, people choose as-if they overweight small probabilities as reflected in Prospect Theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). In contrast, people decide as-if they underweight small probabilities if they acquire risk information sequentially by sampling (Hertwig et al., 2004). The difference in choice patterns between decisions from description and experience has been labeled the *description-experience gap* (DE gap). The difference between decisions from description and experience can in fact be traced to Knight (1921). He suggested that there are (a) a priori probabilities where events result from precisely known stochastic mechanism for which a probability can be assigned; (b) statistical probabilities where the mechanism is not known and probabilities are assessed in an empirical manner; (c) estimates where no probability, neither from a mechanism nor from empirical assessment can be deduced, commonly referred to as uncertainty (Hau, Pleskac, & Hertwig, 2010). The range from apriori probabilities to cases of real uncertainty can further be understood as representing a continuum of

degrees of uncertainty with corresponding external representations, from typically used summary descriptions in terms of probabilities, to the experiences of larger or smaller samples observed in daily life situations (Rakow & Newell, 2010).

In lotteries, outcomes depend on environmental risk alone, whereas outcomes in social dilemmas also depend on the choices of other individuals. Stochastic social dilemmas thus combine social uncertainty and environmental risk. Yet our understanding of cooperation in stochastic environments is currently limited to situations in which environmental risk is described by outcomes and probabilities (Bereby-Meyer & Roth, 2006; Gangadharan & Nemes, 2009; Kunreuther, Gabriel Silvasi, Eric T. Bradlow, & Dylan Small, 2009; Levati, Morone, & Fiore, 2009; Gong, Baron, & Kunreuther, 2009). We argue that real-world risky choices often involve experiencing the outcomes and probabilities of choices rather than receiving their summary statistics. Therefore, examining how risk presentation influences people's decisions is critical to understand how and when people cooperate in risky environments.

There is one important presupposition: risk presentation can influence cooperation only if people are responsive to differences in environmental risk. In lotteries, people's decisions have been found to vary with *different levels of risk*, i.e. for different combinations of outcomes and probabilities while keeping the expected value constant. Analogously, one can describe a stochastic social dilemma by the expected payoffs of cooperation. In a one-shot prisoner's dilemma, people not only cooperate but also respond to different outcomes (Guyer & Rapoport, 1972). Extending this finding to a stochastic setting, we do not only vary the outcomes but also the probability with which they occur. The second goal of this study is thus to establish whether and how different levels of risk affect behavior in one-shot social dilemmas with the same expected payoffs.

We evaluate how different levels of risk affect behavior in a stochastic social dilemma by using a novel approach based on Prospect Theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) that ranks decision situations according to their Prospect Theory Values. Because cooperation rates in social dilemmas are frequently below 50%, majority predictions cannot be used. In contrast, the ranking allows relative predictions among the decisions situations which can be applied to social dilemmas and lotteries in the same way. Prospect Theory, which has been originally proposed to describe risky choice, has recently been applied to a stochastic social dilemma: choices in a stochastic public

good's game show that the degree of cooperation depends on whether the probability of a loss does or does not exist (Iturbe-Ormaetxe et al., 2011; McCarter, Rockmann, & Northcraft, 2010). The authors of both studies interpret the finding as loss aversion, a core element of Prospect Theory. Prospect Theory keeps the original weight-and-add framework of Expected Value Theory and models people's choices as-if they integrate all outcomes and probabilities, which in addition are non-linearly transformed. Compared to their expected value, small probabilities are thus overweighted and large probabilities are underweighted in Prospect Theory.

A different tradition of modeling human behavior is represented by process theories which aim to capture the decision process that people actually employ.[2] For instance, the priority heuristic (Brandstätter, Gigerenzer, & Hertwig, 2006) represents a lexicographic strategy according to which people sequentially evaluate cues instead of integrating all cues.[3] The authors show that it can predict empirical data better than modifications of Expected Utility Theory, such as Cumulative Prospect. Ideally, competing theories from both traditions should be used. Yet in order to make predictions for the decision situations in games and lotteries, the models would first need to be adapted to either predict a ranking or the proportions of risky choices in social dilemmas and lotteries in the same way.

Like other types of choices, cooperation is a function of the match between decision processes and the decision-making environment, or what has been labeled ecological rationality (Simon, 1956; Smith, 2008; Todd, Gigerenzer, & the ABC Research Group, 2012). Besides social uncertainty, which has been studied extensively, the levels of environmental risk and uncertainty are critical components of real-world environments that researchers are only recently beginning to appreciate. For instance, Bereby-Meyer and Roth (2006) investigate the role of learning and find that cooperation unravels slower in a stochastic Prisoner's Dilemma than in a deterministic one. Gong, Baron, and Kunreuther (2009) compare cooperation between groups and individuals, and find that groups cooperate more than individuals in a stochastic Prisoner's Dilemma compared to a deterministic setting. Kunreuther, Silvasi, Bradlow, and Small (2009) further show that people are likely

[2] Pioneering work in risky choice has been done by Rubinstein (1988) and Leland (1994).
[3] The priority heuristic contains the following steps: 1. Priority rule: Go through the attributes of lotteries in the order of minimum gain, probability of minimum gain, maximum gain. 2. Stopping rule: stop information search if the minimum gains differ by one tenth or more of the maximum gain, otherwise stop search if probabilities of minimum gain differ by .1 or more. 3. Decision rule: Choose the lottery that is more attractive in the attribute which stopped search.

to cooperate in a deterministic than in a stochastic Prisoner's Dilemma because in the latter they also respond to whether or not they suffered a loss in the last round. None of the studies, however, addresses how differences in risky environments and the way risk is presented affects cooperation.

Experiment

The goal of the study is to investigate how risk presentation and different levels of environmental risk affect cooperation in a social dilemma. Even if the outcomes of cooperation also depend on the action of others, the environmental risk affects all who cooperate equally. We thus expect both aspects to influence cooperation in risky environments in the same way as lottery choices with environmental risk alone. We present the detailed hypotheses after the Methods section to facilitate understanding.

We used a 2 x 2 between-subjects design in which we manipulated risk presentation (description vs. experience) and choice situation (social dilemma vs. lottery). In the *description* condition, subjects received information about how environmental risk influenced outcomes in a *social dilemma* as a probability statement, whereas in the *experience* condition participants sampled to infer the probabilities. To control whether the values and probabilities chosen to implement environmental risk replicated the DE gap, two further groups made decisions in *lotteries*, again either from description or experience. The environmental risk was identical between lotteries and games. To investigate how different levels of risk affect behavior in one-shot social dilemmas, we varied probabilities and outcomes within-subjects while keeping the expected outcomes constant.

Methods

Environmental risk in social dilemmas and lotteries. For the *social dilemma* conditions, we used a stochastic 2-person public goods game (PG) with binary choices. For each choice, participants receive an endowment e (10 €) which they could contribute to a joint project with a randomly matched partner or keep for themselves. Contributions were

multiplied by a value (*msr*) and shared equally between both pair members.[4] Denoting *i*'s contribution by c_i, where $c_i \in \{0, e\}$ and $i = 1, 2$, *i*'s payoff is given by

$$\pi_i = e - c_i + \frac{msr}{2}(c_1 + c_2). \qquad (1)$$

We impose $msr \in \{1, 2\}$. An $msr > 1$ made it socially optimal to contribute, whereas an $msr < 2$ rendered free-riding the dominant strategy for a selfish person, thus creating a social dilemma.

We manipulated environmental risk by assigning the *msr* to one of two possible values, representing either a good or a bad event, with a certain probability. In case the bad event occured (with probability *p*), contributions were multiplied by an $msr < 1$, decreasing the value of the public good. When the good event occured, contributions were multiplied by an $msr > 1$, increasing the value of the contributions. The environmental risk only affected what was invested. Cooperation thus represents the risky and non-cooperation the sure option. We chose the two potential *msr*-values and corresponding probabilities such that the expected *msr*, *E[msr]*, across good and bad event always yielded a social dilemma with $1 < E[msr] < 2$.

Table 1 illustrates the eight decision situations employed. Situations 1 to 4 contained *rare* ($p < .25$) bad events, analogous to the DE gap studies with lotteries (e.g., Hertwig et al., 2004). Situations 5 and 6 contained more *common* ($p > .25$) bad events to test whether the DE gap extends beyond rare events as found by Ludvig and Spetch (2011a). We use two different expected *msr*, 1.2 and 1.4, to check the robustness of the results. Situations 1 – 6 were designed to extend the findings from the DE gap studies in risky choice to social dilemmas. At the same time, keeping the expected *msr* constant across different combinations of probabilities and potential returns allows us to test whether different levels of environmental risk affect choices in the PG in the same way as they affect choices in lotteries.

Decision situation 7 and 8 explored extreme conditions of a social dilemma and provided a further control of participants' understanding of the incentives. In situation 7, the *E[msr]* equaled 1.1, which made it much less attractive to cooperate compared to situa-

[4] Public goods games conventionally use the marginal per capita return (*mpcr*), i.e., what each participant earns from one unit of investment to the public good. The *msr* is equal to *mpcr**n, where n is the number of interacting subjects. We used the *msr* because it is more easily understood that values below 1 decrease the value of the public good compared to what has been contributed, while values above 1 increase its value. Using the *msr* thus makes it easier for participants to understand the impact of the *msr* for their earnings.

tions 1 – 6. In contrast to the other situations, in situations 7 the rare event was the good state of the world. Different from situations 1 to 7, the expected *msr* of 2.1 in situation 8 did not generate a social dilemma and made it individually and socially optimal to cooperate.

Table 1

Decision Situations in Lotteries and Corresponding Matrices for Stochastic Public Goods Dilemmas

Nr.	Lotteries		Public goods Dilemmas											
	Risky option	Expected msr		Expected msr				Good state				Bad state			
				C		NC		C		NC		C		NC	
Rare events															
1	1.30, 92% 0, 8%	1.2	**C**	1.2	1.2	0.6	1.6	1.3	1.3	0.7	1.7	0.0	0.0	0.0	1.0
			NC	1.6	0.6	1.0	1.0	1.7	0.7	1.0	1.0	1.0	0.0	1.0	1.0
2	1.45, 80% 0, 20%	1.2	**C**	1.2	1.2	0.6	1.6	1.5	1.5	0.7	1.7	0.0	0.0	0.0	1.0
			NC	1.6	0.6	1.0	1.0	1.7	0.7	1.0	1.0	1.0	0.0	1.0	1.0
3	1.55, 92% 0, 8%	1.4	**C**	1.4	1.4	0.7	1.7	1.6	1.6	0.8	1.8	0.0	0.0	0.0	1.0
			NC	1.7	0.7	1.0	1.0	1.8	0.8	1.0	1.0	1.0	0.0	1.0	1.0
4	1.80, 80% 0, 20%	1.4	**C**	1.4	1.4	0.7	1.7	1.8	1.8	0.9	1.9	0.0	0.0	0.0	1.0
			NI	1.7	0.7	1.0	1.0	1.9	0.9	1.0	1.0	1.0	0.0	1.0	1.0
Two common events															
5	1.80, 64% 0.20, 36%	1.2	**C**	1.2	1.2	0.6	1.6	1.8	1.8	0.9	1.9	0.2	0.2	0.1	1.1
			NC	1.6	0.6	1.0	1.0	1.9	0.9	1.0	1.0	1.1	0.1	1.0	1.0
6	1.95, 56% 0.70, 44%	1.4	**C**	1.4	1.4	0.7	1.7	2.0	2.0	1.0	2.0	0.7	0.7	0.4	1.4
			NC	1.7	0.7	1.0	1.0	2.0	1.0	1.0	1.0	1.4	0.4	1.0	1.0
Extreme expected msr															
7	0.75, 88% 3.50, 12%	1.1	**C**	1.1	1.1	0.5	1.5	3.5	3.5	1.8	2.8	0.8	0.8	0.4	1.4
			NC	1.5	0.5	1.0	1.0	2.8	1.8	1.0	1.0	1.4	0.4	1.0	1.0
8	2.20, 96% 0.30, 4%	2.1	**C**	2.1	2.1	1.1	2.1	2.2	2.2	1.1	2.1	0.3	0.3	0.2	1.2
			NC	2.1	1.1	1.0	1.0	2.1	1.1	1.0	1.0	1.2	0.2	1.0	1.0

Note. For each decision situation, the columns for "Lotteries" show the two *msr*, followed by their probability of occurrence and the expected msr. The columns "Public goods dilemmas" show the game matrices based on the expected *msr*, as well the matrices based on each possible *msr*. C stands for contributing, NC for not contributing. In each cell, the left (right) number denotes the payoff of the row (column) player.

In most studies on the DE gap, the risky option has an expected value that is only marginally higher than the sure option.[5] To avoid floor effects in the social dilemma, we used relatively large expected *msr.* This should provide strong incentives to cooperate in the PG, but results in a larger difference between the expected *msr*-value of the sure option and the risky option. To control whether the parameters we chose for implementing environmental risk nevertheless replicated the DE gap in more standard settings, we ran the same choices as lotteries with identical environmental risks. In the lottery conditions, participants also received an endowment *e* and had to decide whether to invest it in a risky option. The risky option in each lottery used the same two possible *msr* with the same probabilities as the corresponding PG. Yet, while the payoffs in the games also depended on the action of another person, the payoffs in the lotteries only depended on the realized state of the world. The lotteries strip the strategic component away but retain the stochastic component that stems from the environment. Consider decision situation 1 as a lottery: if a participant invests her endowment, she can earn either 1.30**e* with 92 % probability or 0**e* with 8 % probability; alternatively, if she does not invest, she earns *e* for sure. If decision situation 1 is a social dilemma and the good (bad) state is realized, each receives 1.30**e* (0**e*) if both pair members invest and *e* if both do not invest. We randomized the order of decision situations in games as well as in lotteries, and participants received no feedback about the realized *msr* (or the decision of the other group member) after each decision.

Decisions from Description vs. Decision from Experience. In the *description* conditions, participants received information about environmental risk as a summary statement about probabilities and associated *mrs*-values before they made their decision. In the *experience* conditions, participants sampled the distribution of *mrs*-values by drawing 25 cards from a deck. We used a matched-sampling design based on Ungemach, Chater and Stewart (2009), where people were forced to view a representative sample of the underlying distributions of outcomes. Each card contained a number corresponding to one of the two possible *msr*. For example, in situation 1 the deck had 2 cards with the *msr* 0 and 23 cards with the *msr* 1.30. The sequence of cards was randomized for each participant, yet

[5] See, for instance, the set of lotteries used by Hertwig et al. (2004a) which have also been used widely by others. Here, the risky option is only about 1.07 times higher than the safe option. However, Hau et al. (2010) employ a lottery with an EV for the risky option that is 1.26 times larger than the sure option and find a DE gap if participants draw a larger number of samples.

the two possible *msr* and their frequencies matched exactly the objective probabilities given in the *description* condition. Thus, a sampling error could not cause any differences observed between the two conditions.

Further Measures. In the *experience* conditions, we additionally collected time stamps that allowed us to evaluate how long participants viewed a certain card and whether this influenced their decision. To check the accuracy of risk estimates, we also asked participants after the last round how often they saw the two sampled *msr*-values.

In the *description* conditions, participants were asked to translate the probability statement of the last round into a frequency statement, so that we could control whether participants had accurately understood the probability information presented to them.

In the *social dilemma* conditions, participants also faced two deterministic PGs with an *msr* of 1.2 and 1.4 (randomized order) after the stochastic situations. This allowed us to investigate how cooperation varies if the stochastic component is removed, since the deterministic games matched the expected *msr* of the stochastic PGs in situations 1, 2, and 5 ($E[mrs] = 1.2$) as well as 3, 4, and 6 ($E[mrs] = 1.4$).

At the end of the experimental session, participants indicated which reason best explained their decision to invest and not invest into the stochastic PGs by answering to a questionnaire adapted from McCarter, Rockmann, and Northcraft (2010a). Participants could select one of the following six reasons (see Appendix A2 for questionnaire): the probability of the *mrs* were (not) sufficiently high, the values of the *mrs* were (not) sufficiently high, conditional cooperation, social uncertainty, greed/opportunism, moral values, or none of these. A section on demographics concluded the experiment.

Participants and Procedure. We recruited 128 students from Jena, Germany, using the ORSEE software (Greiner, 2004), and randomly assigned them to one of four sessions. In the social dilemma conditions, participants had to pass control questions to ensure that they understood the impact of environmental risk and of the other person's choice on their payoffs (see Appendix A3 for experimental instructions). Participants were generally relatively unfamiliar with laboratory experiments.

All tasks were completed anonymously, employing a perfect stranger design in the social dilemma conditions. The lottery sessions lasted 60 minutes, and the game sessions

75 minutes. At the end, one decision situation was randomly chosen to determine the payoff. Participants earned 14 € on average. The experiment was conducted at the laboratory of the Max Planck Institute of Economics in Jena using z-Tree (Fischbacher, 2007).

Hypotheses

Risk sensitivity in social dilemmas and lotteries. How do different levels of risk affect choices in stochastic PGs compared to choices in lotteries? We examined this question by focusing on decisions from description and testing the predictions of Prospect Theory (Tversky & Kahneman, 1992). Using a separate value and weighting function, Prospect Theory transforms the expected outcomes of a lottery into Prospect Theory Values (PTVs), analogous to expected values. When comparing the PTV of a lottery's risky option with a sure option (always 1 in our case), the conventional prediction is that the risky (sure) option is picked by the majority if the PTV is larger (smaller) than the one of the sure option. However, investment rates in the PG are expected to be lower than in lotteries due to a second source of uncertainty that stems from the other person. Thus, one cannot use predictions for the majority choice based on PTVs relying on environmental risk alone. However, the PTVs also allow a ranking of the 8 decision situations. The ranking can be interpreted as predicting whether the proportion of risky choices in each decision situation should be higher or lower relative to the other situations. Such a ranking can be applied to both lotteries and stochastic PGs in the description condition. Table 2 lists the PTVs for the eight decision situations of this experiment based on the parameters used by Tversky and Kahneman (1992). From the PTVs, two predictions follow for PGs and lotteries with the same expected *msr*:

(1a) Situations 1 and 3 (bad event occurs with 8%) will lead to a higher number of risky choices than situations 2 and 4 (where the bad event occurs with 20%).

(1b) Situation 5 and 6, where the bad event is more common, will lead to more risky choices than situations 1 and 2 or 3 and 4.

Table 2

Prospect Theory Values for each Decision Situation and Predicted Ranking of Proportions of Risky Choices

	Risky Option	Expected msr	Prospect Theory	
			PTVs	Ranking
Rare event				
1	1.30, 92% 0, 8%	1.2	0.93	II
2	1.45, 80% 0, 20%	1.2	0.84	III
3	1.55, 92% 0, 8%	1.4	1.09	B
4	1.80, 80% 0, 20%	1.4	1.02	C
Two common events				
5	1.80, 64% 0.20, 36%	1.2	0.96	I
6	1.95, 56% 0.70, 44%	1.4	1.21	A
Extreme expected msr				
7	0.75, 88% 3.50, 12%	1.1	1.23	
8	2.20, 96% 0.30, 4%	2.1	1.70	

Note. Each row describes one decision, by first listing the two *msr* followed by their probability of occurrence and the expected *mssr*. The value of the sure option is always 1. The ranking is done for decision situations with the same expected *msr*, denoted by roman numbers I – III for *E[msr]* = 1.2, and by letters A-C for *E[msr]* = 1.4.

Decisions from Description and from Experience. Using lotteries, studies found that experienced small probabilities appear to be underweighted in choices compared to small probabilities that are learned based on a summary description (Rakow & Newell, 2010). Extending this choice pattern to social dilemmas leads to the following hypothesis for stochastic PGs and lotteries:

> (2) The risky option will be chosen more frequently in the experience condition than in the description condition if the bad event is less likely (situations 1 – 6 and 8), whereas this pattern should reverse for situation 7, in which the good event is less likely.

Results

Risk Sensitivity in Social Dilemmas and Lotteries

We would not expect risk presentation to matter unless people are sensitive to different levels of risk in games as they are in lotteries. For the results of hypothesis 1a and 1b, we focus on data from the *description* conditions for decision situations 1 to 6. Figure 1 illustrates the behavior in lotteries and stochastic PGs separately for the expected *msr* of 1.2 and 1.4. For comparison, it includes the rate of investment in the deterministic PG.

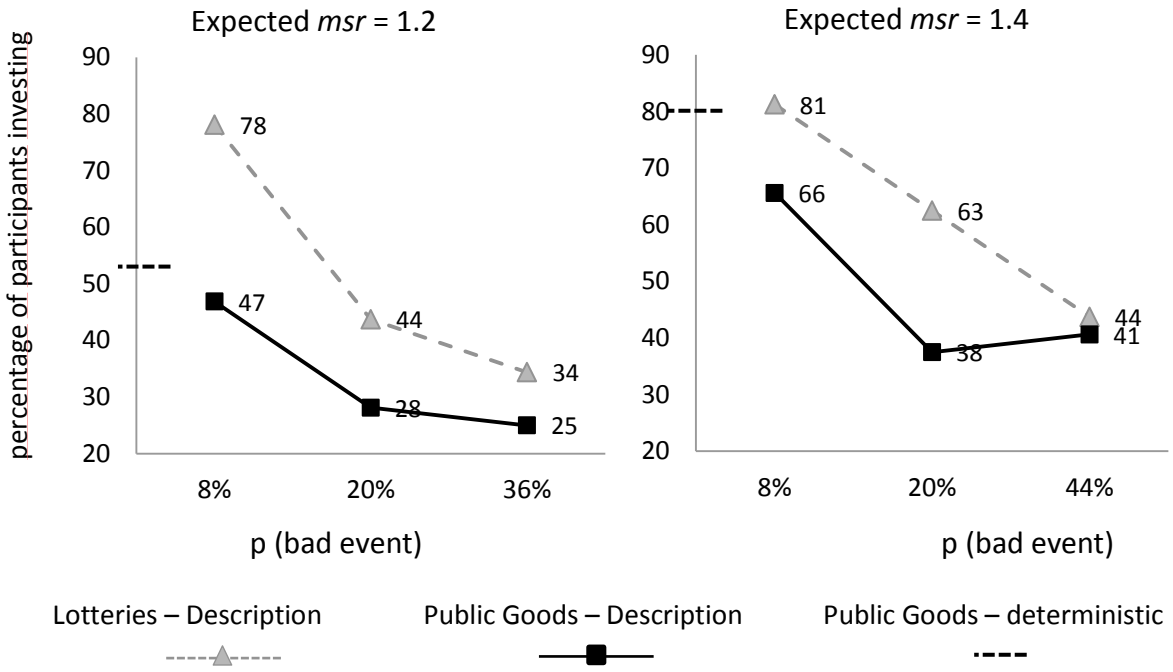


Figure 1. Investment rates in lotteries and Public Good games for the description condition, as well as for the deterministtic Public goods games. The x-axis displays the probability of the bad (less likely) *msr* of a decision situation and the y-axis the percentage of participants who choose to invest into the risky option. The line crossing the y-axis shows the investment rate in deterministic games.

When comparing decision situations with an *E[msr]* = 1.2 and *E[msr]* = 1.4 (Figure 1, left and right panel), cooperation increases with the expected *msr*. The deterministic PGs yield a similar pattern: the rate of cooperation is 53% when *msr* = 1.2 and, 81% when *msr* = 1.4 ($\chi^2(1) = 5.74$, p = .02). In the stochastic PGs, the average rate of cooperation is 33% when *E[msr]* = 1.2 and 48% when *E[msr]* = 1.4 ($\chi^2(1) = 4.23$, p =.04). Thus, differences in

expected *msr* affect behavior even though the social dilemma is maintained and the dominant strategy for a person is not to cooperate. This replicates Guyer & Rapoport (1972) findings and extends it to a stochastic setting. But, besides being sensitive to different expected outcomes, do people react to different levels of risk for constant expected outcomes*?*

To address this question, we pool our data across situations with expected *msr*-values of 1.2 and 1.4 to obtain more reliable results.[6] The mean cooperation rate is 1.7 times higher in situations where the bad event occurs with 8% than in situations where the bad event is common ($\chi^2(1) = 7.12$, p =.01). Thus, changes in the stochastic environment have a large impact on cooperation. Note that the difference in cooperation between deterministic and stochastic PG with an 8% chance of a bad event is only 10.5% and not significant ($\chi^2(1) = 1.62$, p =.20).

To investigate hypothesis 1a – that situations with 8% receive more investment than situations with 20% –, one can also rely on the pooled data across the *E[msr]* of 1.2 and 1.4 because the rankings of PTVs are identical for both (Table 2). The rate of investment in situations with a probability of 8% compared to 20% sharply drops both for stochastic PGs (from 56% to 33%, $\chi^2(1) = 7.17$, p =.01) and lotteries (from 80% to 53%, $\chi^2(1) = 10.12$, p <.001). Paralleling each other, stochastic PGs and lotteries thus are in line with prediction 1a based on Prospect Theory.

For prediction 1b – that the rate of investment is highest with a common event –, Figure 1 also suggests a decline in cooperation between situations with a probability of 20% and those with two common events. Statistically, however, there is no difference between these two situations, neither for the stochastic PGs (pooled across *E[msr]* of 1.2 and 1.4 again, the investment rate is constant at 33%, $\chi^2(1) = 0.00$, p = 1.00), nor for lotteries (the investment rate declines from 53% to 39%, $\chi^2(1) = 2.55$, p =.11). Hence, hypothesis 1b based on Prospect Theory – that the rate of investment is highest with a common event – is neither met in stochastic PGs nor in lotteries.

In summary, we find that different levels of environmental risk both influence choice in the PGs for decisions from description and result in similar behavior in stochastic

[6] Earlier studies report considerable noise in lotteries (Harless & Camerer, 1994; Hey & Orme, 1994; for a summary see Wakker, 2010). Employing a repeated within-subject design, they find that participants are inconsistent in about one out of four choices.

PGs and lotteries. Though the data confirm the predictions of Prospect Theory for hypothesis 1a, we did not obtain support for hypothesis 1b for either PGs or lotteries.[7]

Decisions from Description and from Experience

In this section, we address how risk presentation affects behavior in games and whether it does so in the same way as in lotteries (hypothesis 2). Afterwards, we explore the reasoning processes in stochastic PGs in the experience and description condition in comparison to those in lotteries.

Is there a DE gap in lotteries and games? We initially focus on pooled data from the eight decision situations to start with more reliable results. Hypothesis 2 is directional and states that, except for situation 7, participants should choose the risky option more often in the experience condition. To test this hypothesis, we subtracted the percentage of people contributing in the experience condition from those in the description condition, except for situation 7 where we do the opposite. Figure 2 illustrates the DE gap for lotteries and stochastic PGs. The results show a positive gap for lotteries (χ^2 (1) = 8.24, p =.003), with a mean difference between experience and description of 12% (SD = 10%).

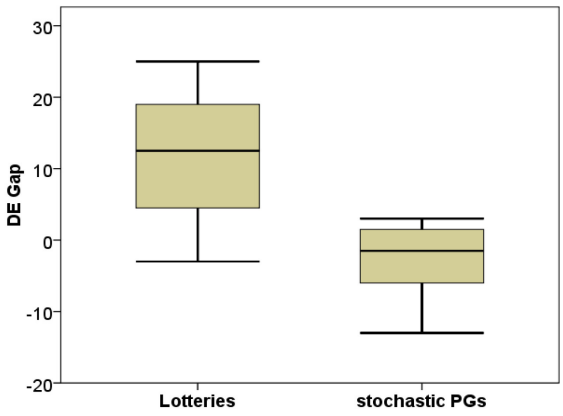


Figure 2. Box plots of the DE gap pooled across eight situations in lotteries and games. The box plot displays the minimum, first quartile, median, third quartile, and maximum of the DE gap.

[7] For comparison, the constant-relative-risk-aversion utility of the decision situation can be calculated for instance by using the parameterization in Goeree, Holt, and Palfrey (2003). However, the predictions are also not met. The calculations for instance suggest that participants should choose the risky option more often in decision situation 6 than in situation 3 which is neither met in games nor lotteries.

Table 3 lists the percentage of people investing in experience and description separately for all eight decisions situations in lotteries and stochastic PGs. For lotteries, the predicted difference between the experience and description condition is observed in all situations (including the reversal for situation 7) – except for lottery 8. This lottery shows a ceiling effect because the expected outcome is twice as high as the sure option, so that in both conditions all participants but one invested.

Table 3

Percentage of Subjects Investing in Public Good games and Lotteries and Differences between Description and Experience Condition.

Decision Situations				Stochastic PG			Lotteries		
#	Risky Option	*E[msr]*	PTV	Desc	Exp	Difference between description and experience conditions	Desc	Exp	Difference between description and experience conditions
One rare event									
1	1.30, 0.92 0, 0.08	1.2	0.93	47	44	-3 ($\chi^2(1) = 0.06$, p = .80)	78	81	+3 ($\chi^2(1) = 0.10$, p = .76)
2	1.45, 0.8 0, 0.2	1.2	0.84	28	28	0 ($\chi^2(1) = 0.00$, p = 1.00)	44	69	+25 ($\chi^2(1) = 4.06$, p = .04)
3	1.55, 0.92 0, 0.08	1.4	1.09	66	56	-9 ($\chi^2(1) = 0.59$, p = .44)	81	88	+6 ($\chi^2(1) = 0.47$, p = .49)
4	1.80, 0.8 0, 0.2	1.4	1.02	38	38	0 ($\chi^2(1) = 0.00$, p = 1.00)	63	78	+16 ($\chi^2(1) = 1.87$, p = .17)
Mean 1 – 4				*45*	*41*	*-3 ($\chi^2(1) = 0.26$, p = .61)*	*66*	*79*	*+13 ($\chi^2(1) = 5.03$, p = .03)*
Two common events									
5	1.80, 0.64 0.20, 0.36	1.2	0.96	25	28	3 ($\chi^2(1) = 0.08$, p = .77)	34	44	+9 ($\chi^2(1) = 0.59$, p = .44)
6	1.95, 0.56 0.70, 0.44	1.4	1.21	41	28	-13 ($\chi^2(1) = 1.11$, p = .29)	44	59	+16 ($\chi^2(1) = 1.56$, p = .21)
Mean 5 & 6				*33*	*28*	*-5 ($\chi^2(1) = 0.33$, p = 0.57)*	*39*	*52*	*+13 ($\chi^2(1) = 2.02$, p = .16)*
Extreme *msr*									
7	0.75, 0.88 3.50, 0.12	1.1	1.23	19	16	-3 ($\chi^2(1) = 0.11$, p = .74)	38	16	-22 ($\chi^2(1) = 3.92$, p = .05)
8	2.20, 0.96 0.30, 0.04	2.1	1.70	91	88	-3 ($\chi^2(1) = 0.16$, p = .69)	100	97	-3 (p = .50, Fisher's exact test)

Note. In situation 8, Fisher's exact test is applied, due to a cell having only a count of 1. Per condition n = 32

Averaging across lotteries 1-4, which contain a rare event, shows a DE gap of 13% (Table 3). The same DE gap (13%) occurs with lotteries containing a more common bad event (situation 5 and 6, Table 3). The results replicate Ludvig and Spetch (2011), who find the DE gap also in situations with common events. Overall, responses to decisions

from description and experience differed in lotteries, in line with previous findings. Thus, the parameters we chose for environmental risk replicate the DE gap found in the risky choice literature.

Given that the parameters replicate the DE gap in lotteries, and given the previous result that people's decisions in games were similarly sensitive to differences in risk as in lotteries, we expected the risk presentation format to influence cooperation as well. The behavior in the stochastic PGs, however, does not match the behavior in lotteries in this respect: the DE gap completely disappears in games ($\chi^2(1) = 0.38$, $p = .30$). The mean difference between experience and description in the stochastic PG is -3% (SD = 6%).

The stochastic PGs stand in stark contrast to the results in the lotteries. In games, 6 out of 8 decision situations show no or only minimal gaps (Table 3). Experience and description conditions do not differ for any of the decision situations. In fact, situation 7, which resembles the situations used by Hertwig et al. (2004) and Ungemach et al. (2009) more closely, shows a strong DE gap in lotteries, but the gap is completely absent in games.

Why is there a DE gap in lotteries but not in games? In the following, we explore the reasoning processes in PGs and lotteries which provide hints to why risk presentation affects lotteries but not stochastic PGs.

One possible explanation underlying this pattern is that participants spend different amounts of time sampling in lotteries and games (Figure 3), which may indicate different search processes. In lotteries, the average participant spent more time viewing the rare event (M = 1.04 seconds, SD = 0.67) compared to the more frequent event of each decision situation (M = 0.68 seconds, SD = 0.25, $t(31) = 3.58$, $p = .001$). Similarly, for the games, the average participant viewed the rare event (M = 0.53 seconds, SD = 0.23) longer than the more frequent event (M = 0.44 seconds, SD = 0.14, $t(31) = 3.72$, $p = .001$). However, overall participants spent more time sampling than in games for both rare events ($t(38.24) = 3.99$, $p = .001$) and frequent events ($t(47.78) = 4.87$, $p = .001$). These differences in sampling times thus provide evidence for potentially different search processes by which people appear to pay less attention to the actually observed probabilities in games compared to lotteries with identical environmental risk.

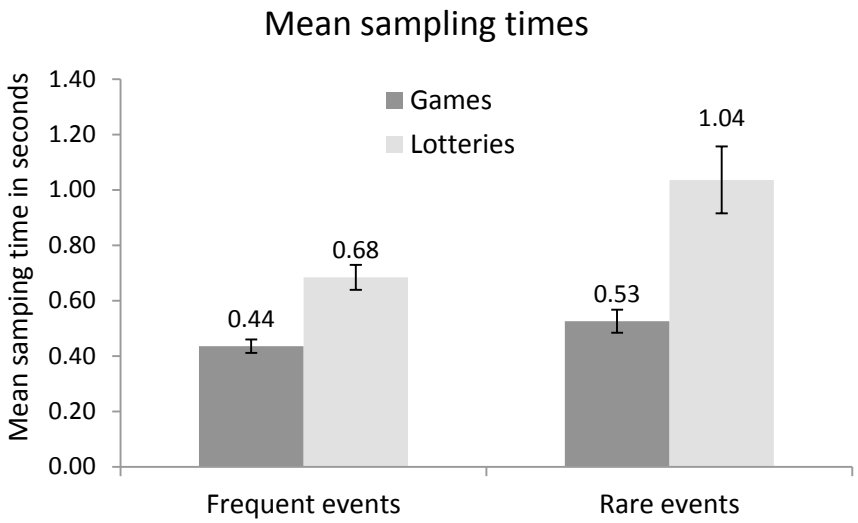


Figure 3. Mean time in seconds spent per draw for rare and more frequent events of the decision situations in lotteries and games. Displayed means were calculated across the mean sampling time of each participant across decision situations, separately for rare and more frequent events. Error bars indicate one standard error above and below the mean.

To control for the accuracy of risk perception, participants in the experience conditions stated the frequency of the two outcomes in the last situation after they had decided. The actual distribution of outcomes that participants experienced correlates with the stated frequencies for lotteries ($r_S = 0.72$, $p < .001$) yet to a lesser extent for stochastic PGs ($r_S = 0.43$, $p < .01$). In both conditions, participants were calibrated to the actual probabilities and did not underestimate but, if anything, overestimated the probability of rare events (see Figure A1 in the Appendix).

Some researchers have suggested that DE gap may be driven by the larger influence of recent events in decisions from experience. While some studies have found a recency effect in decisions from experiences (Hertwig et al., 2004, Rakow, Demes, & Newell, 2008), others have not (Ungemach et al., 2009, Hau, Pleskac, Kiefer, & Hertwig, 2008). To test for a recency effect, we divided the 25 samples participants draw before each decision into two sets: from 1 to 12 (initial) and from 13 to 25 (latter). Then we computed the expected *msr* from the initial samples, $E[msr]_{1\text{-}12}$, and from the latter samples, $E[msr]_{13\text{-}25}$. Finally, we compare the number of risky choices made when $E[msr]_{13\text{-}25} > E[msr]_{1\text{-}12}$ to the number of risky choices made when $E[msr]_{13\text{-}25} < E[msr]_{1\text{-}12}$. When the expected *msr* of the latter, more recent sample ($E[msr]_{13\text{-}25}$) was larger, we find a higher number of risky

choices in lotteries ($\chi^2(1) = 3.77$, p =.04) but not in games ($\chi^2(1) = 0.30$, p =.34). This also suggests that the actual observed probabilities may play a less important role in games than in lotteries (see Table A1 in the Appendix).

Finally, for the stochastic PG in description and experience, participants indicated their most important reasons for cooperating as well as not cooperating which are shown in Figure 4. This resulted in two reasons per participant (see Table A2 in the Appendix). Aggregating across both statements, probabilities influenced cooperation decisions in the description condition for 59% of the participants, compared to 39% in the experience condition. In the experience condition, participants rather emphasized both the value of the *msr* they could obtain (20% in experience compared to 3% in description) and their expectation of whether the other will (not) cooperate, i.e. conditional cooperation (20% in experience compared to 11% in description). This indicates that the importance of the probabilities for the decisions is further reduced in the stochastic PG in experience.

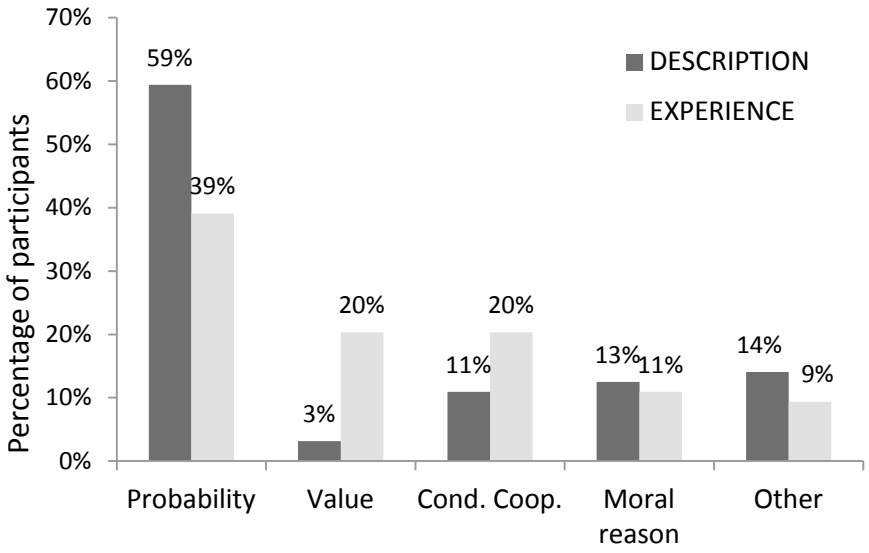


Figure 4. Most important reasons for contributing and not contributing in stochastic Public Goods games, separately for description and experience condition.

In summary, participants sampled more quickly in the stochastic PG in the experience condition than in lotteries, as if they were paying less attention to the observed probabilities. In line with this, subjects' risk perception was less accurate in games than in lotteries, and recency – a potential cause of the DE gap – did not play a role in games, whereas we did find a recency effect in lotteries. The questionnaire also highlighted that probabilities were less important in the PG in experience than the size of the values and

beliefs about others' behavior. This provides converging evidence that, as the probabilities of the risky option lose importance in the games, the DE gap washes out.

Discussion

People often cooperate in social dilemmas. We examined how critical aspects of the stochastic environment shape cooperation. First, different levels of environmental risk influence cooperation. Investments decisions in the stochastic PGs match those observed in lotteries, with people preferring an 8% chance of a bad event to a 20% chance for constant expected payoffs. Second, the *msr*-values and probabilities chosen to implement environmental risk replicate the DE gap within individual risky choices in lotteries. That is, people choose the risky option more often when experiencing the risky outcomes compared to when receiving summary descriptions. Our key finding is that, nevertheless, risk presentation matters in lotteries but not in games: no DE gap was observed for the social dilemmas. Process data and subjects self-reported reasons for cooperation suggest that the absences of a DE gap in games may result from a decision process that emphasizes the size of the outcomes and participants' expectations about others' behavior over outcome probabilities.

From a Prospect Theory perspective, choices in decisions from experience have been described as inversing the non-linear weighting of probabilities found in decisions from description, which leads to the gap in choices between both conditions (Hertwig et al. 2004; Ludvig & Spetch, 2011; Ungemach et al., 2009). In our study, however, Prospect Theory could not even account for the data in the description condition for either lotteries or games. One possible reason could be that the values we have chosen for the stochastic environment invites violations of stochastic dominance which Prospect Theory cannot account for (Birnbaum, Coffey, Mellers, & Weiss, 1992). In particular, Wakker (2010) points out that 'zero' outcomes are liable to result in behavior not in line with Prospect Theory. This shows the limits of introducing Prospect Theory to a stochastic social dilemma (Iturbe-Ormaetxe, Ponti, Tomás, & Ubeda, 2011; McCarter, Rockmann, & Northcraft, 2010)

In our view, to include environmental risk and decisions from experience into the study of cooperation invites more realism into the laboratory. This study is only a small step to build on insights from research on risky choice for decision situations which combine environmental risk and social uncertainty. An important goal for future research is to

model the actual decision processes, by relying more on process and sampling data. Instead of focusing on choices alone, such models provide promising alternative starting points to weight-and-add models in the tradition of Expected Value Theory (Gigerenzer, Todd, & The ABC Research Group, 1999; Simon, 1956; Smith, 2003) such as Prospect Theory, which in our study could not account for the data in the description condition for either lotteries or games. In fact, the basic message of our study – that outcomes feature more strongly than probabilities in some environments but not in others – has been pointed out elsewhere in the risky choice literature (e.g., Sunstein, 2003).

In complex interactive environments, it seems rather likely that non-compensatory decision making emerges. For instance, a lexicographic strategy like the Priority Heuristic (Brandstätter et al., 2006), outlines a sequential decision process which considers outcomes as a first and probabilities only as a second step if no decisions could be made based on the first step. In a similar fashion, other strategies that do not trade-off reasons may be valuable to model search and decisions processes in situations that combine environmental risk and social uncertainty – and thus also include expectations about others and further social reasons besides mere outcomes and probabilities.

Acknowledgments

This research was funded by a research grant of the Arete Initiative at the University of Chicago, with additional support from the Max Planck Society.

Chapter 4

Moral Hindsight: Moral Judgments under Certainty and Uncertainty[1]

Nadine Fleischhut [a] & Gerd Gigerenzer [a]

[a] Max Planck Institute for Human Development, Berlin, Germany

[1] Fleischhut, N., & Gigerenzer, G. (manuscript in preparation). Moral hindsight: Moral judgments under certainty and uncertainty.

Abstract

How do people make moral judgments in dilemmas under uncertainty, as opposed to in the commonly investigated situations where the course of all future events is assumed to be certain? We investigated people's moral judgments in two types of epistemic situations: when it is uncertain whether an action will lead to negative side effects (i.e., in foresight), and when it is already certain that negative side effects did or did not occur (i.e., in hindsight). In order to examine how moral judgments are influenced by what is foreseen, we asked participants to estimate the probability of side effects. Our key result is a hindsight effect in moral judgment. Participants in foresight judged actions to be morally more permissible than did participants in hindsight who knew that negative side effects occurred. Conversely, when hindsight participants knew that no negative side effects occurred, they judged actions to be more permissible than did foresight participants, for whom the occurrence of side effects was uncertain. This result raises concerns about generalizing findings from laboratory studies using moral dilemmas in which everything is certain to real world situations in which nothing is. A hindsight effect was also obtained for probability estimates. Both results align in a way consiststent with a consequentialist approach. However, the correlation between probability estimates and moral judgments was only moderate, and the analysis of people's most important reasons further indicated that not everyone took probabilities into account. Because some judgment processes may be less sensitive to uncertainty than a consequentialist process, this further underlines the importance to study both, the cognitive processes and the epistemic conditions that these processes may be adapted to.

Keywords: moral judgment, moral dilemma, uncertainty, hindsight, outcome bias

Introduction

Despite ongoing efforts, hunger catastrophes still regularly occur all over the world. In fact, in 2011 almost one billion children, men, and women went to bed hungry every night (World Disasters Report, 2011). One suggestion has been to help developing countries to cultivate genetically modified crops, which are more robust to environmental conditions and supposed to improve food production. Proponents of genetically modified crops argue that they could efficiently ensure food availability and reduce malnutrition for hundreds of millions in non-industrialized nations. Opponents, however, warn about the unknown, potentially devastating consequences of such a strategy. According to these critics, the long-term risks include serious consequences for public health, such as severe allergies, as well as the possible destruction of fragile eco-systems and food chains (Herrera-Estrella & Alvarez-Morales, 2001).

Is it morally permissible for the government of a developing country to start cultivating a new genetically modified corn that survives under harsh conditions if the risks for health and environment are uncertain and can only be known in the future? And would your moral judgment be differnet if the corn did in fact have serious consequences for the environment and public health—even if this could not be known for sure when the decision was made?

Uncertainty is inherent in many situations in which moral judgments and decisions are made. Sometimes, some information may be available, such as how likely it is that a drought will occur. In other situations, information may be scarce, for instance about the long-term effects of genetically modified corn. Alternatively, people may simply lack the time to gather more information when decisions have to be made under time pressure. Typically, in most real-world situations, the course of events and the consequences of actions are only certain after the fact. Whereas this epistemic uncertainty has been the focus of much psychological research on inferential and preferential judgment and decision making, it has received surprisingly little attention in the domain of moral judgment (Gigerenzer, 2010).

On the normative side, moral philosophy has been concerned with two questions: what is morally right, and what should we do from a moral point of view? The uncertainty under which moral evaluations have to be made is typically disregarded as irrelevant to the first question and reduced to a mere theoretical problem when it comes to the second.

On the descriptive side, moral psychology has focused on the question of how people make moral judgments and on the cognitive processes underlying these judgments. To this end, the morally relevant consequences of action and inaction are typically presented as being absolutely certain (but see e.g., Baron & Ritov, 2009; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001 for some dilemmas including risk). However, if we are interested in how people make moral judgments in real world situations, as opposed to highly artificial and simplified "small worlds" (Bennis, Medin, & Bartels, 2010; Binmore, 2008; G. Gigerenzer, 2010), in which options are limited and all relevant information is available, we also need to consider the epistemic conditions people find themselves in.

The aim of this article is to empirically investigate moral judgments in dilemma situations in which the possible consequences of an action are uncertain, as compared with moral judgments after the fact, where the actual consequences of an action are already determined. To this end, we examine moral judgments in a more ecologically valid paradigm that mirrors the epistemic conditions under which real world moral judgments are made.

Small-world Moral Dilemmas and Uncertainty

Is it morally permissible to push an innocent person in front of a trolley in order to save the lives of five other people on the tracks? Is it permissible to treat a mother with uterine cancer when the treatment will kill her unborn baby, but not treating her will result in the death of both? And is it allowed to throw some passengers on an overcrowded lifeboat overboard to prevent it from sinking? These situations exemplify the characteristic feature of moral dilemmas: on the one hand, the situation calls for an action to protect the well-being of others; on the other hand, the action itself comes at a moral cost.

Artificial moral dilemmas, such as the famous "trolley dilemma" (Foot, 1967; Unger, 1996), are widely used in both philosophy and psychology. These dilemmas typically simplify the situation by presenting the consequences of action and inaction as certain. Moral philosophers have used such dilemmas as thought experiments to inform normative moral theories by identifying criteria for what is morally right. For instance, according to *consequentialist* moral theories (for an overview, see Scheffler, 1988) an action should be evaluated solely by reference to the value of its consequences (e.g., the number of lives saved given vs. not given the action). From this perspective, the morally right action is the one that maximizes the "good outcomes"—for instance, the number of lives saved by stop-

ping the trolley. However, in some variants of the dilemmas, our moral intuitions forbid actions, such as deliberately pushing a person in front of the trolley, even if this action would maximize the "good outcome," such as the number of lives saved. The intuitions elicited by these simplified dilemmas are used to specify *deontological* rules, rights or duties based on morally important features other than mere consequences.

In the normative context, using moral dilemmas in which all outcomes are certain may appear acceptable at first glance—after all, normative criteria need in some sense to be independent of the contingent epistemic situations people can find themselves in. Yet, this is not as unproblematic as it may seem for two reasons. First, moral theories also aim to provide practical guidance in *actual* moral situations—and applying general normative criteria to a specific situation requires information that ideal-typical dilemmas present as certain, but that are often lacking in real situations due to uncertainty and limited knowledge. For instance, when information about possible outcomes and probabilities is missing, a consequentialist cost–benefit analysis will not be helpful in determining what to do from a normative perspective (Binmore, 2008; Gigerenzer, 2010).[2] Thus, normative theories should not disregard uncertainty if they want to provide practical guidance (Sollie, 2009). Second, the normative answer under certainty is not necessarily the same as under risk and uncertainty, as decision theory has shown within the context of rationality (Gigerenzer & Brighton, 2009). Thus, the use of dilemmas in which everything is certain is not unproblematic for the normative analysis.

In contrast, moral psychologists have focused on the descriptive question of how people make moral judgments and employed such dilemmas in empirical research. As these dilemmas typically pit a consequentialist against a deontological response, they are useful for classifying judgments as being in line with either consequentialist or deontological approaches, as well as for investigating in more detail the factors and cognitive processes that underlie people's judgments (for a review, see Waldmann, Nagel, & Wiegmann, 2012). In these studies, epistemic uncertainty is often deliberately removed—the rationale being that this affords better insights into particular cognitive factors and processes specific to the moral domain (e.g., Hauser, Cushman, Young, Kang-Xing Jin, & Mikhail, 2007). While there is some merit to this argument, this approach nevertheless may

[2] One reaction to the question of how moral criteria can provide guidance under uncertainty is to understand the normative criteria as relative to the evidence available at a given point in time (e.g., Zimmerman, 2010). This approach treats still uncertainty as a theoretical problem, but does not specify any decision procedure.

suffer from the same shortcoming as the normative analysis: the dilemmas used are often not representative of the epistemic conditions under which people actually have to make moral judgments.

In many real world situations, the consequences of the action under consideration are difficult to predict or disagreed upon by experts, may it be due to limited information or to complex interactions. In other situations, people may simply lack the time or resources to gain more information—for instance, when an immediate decision is required. Imagine that an airplane departs from its route for no known reason and is on target for a nuclear power plant. In this case, there will be little time to find out whether this behavior is due to a terrorist attack, potentially requiring the plane to be shot down, or simply results from a navigation error or technical difficulties.

Uncertainty, of course, comes in different variants. According to a seminal classification by Knight (1921), different types of uncertainty can be distinguished. S*ituations of risk* refer to cases in which objective probability information for events can either be deduced a priori or estimated from data. The possible outcomes when rolling an unbiased die can be calculated a priori; the negative side effects of medical treatments can be estimated from data. In *situations of uncertainty*, by contrast, no such probabilities are available.[3] For instance, we cannot yet say how likely it is that genetically modified crops will reduce biodiversity, or that an influenza virus will mutate to a dangerous type. People may still form subjective beliefs about the likelihood of such events, but under uncertainty there is no objective basis by which to judge them (LeRoy & Singell, 1987). Yet even in situations of risk, where the chances of an outcome can in principle be estimated, individuals often lack the time and resources to do so and are thus—from an epistemological point of view—also in a situation of uncertainty.

Within a dilemma situation, various morally relevant aspects may be uncertain. The first aspect is the harm that an action is supposed to prevent. For instance, in a widely used version of the trolley dilemma (Foot, 1967), a train will kill five people if it is not redirected to a different track, where it will kill one person. Yet, in real life, it would be uncertain whether some or all of the five people on the track might save themselves in time,

[3] Knight's distinction has been interpreted differently, e.g. as the distinction between what is measurable vs. unmeasurable, objective vs. subjective probabilities (Friedman, 1976), or insurability vs. uninsurability (e.g., LeRoy & Singell, 1987). Depending on the interpretation, a different explanation is given for the claim that here is no valid basis for classifying instances under uncertainty (see e.g., Langlois & Cosgel, 1993).

even if no action were taken. Second, the undesired side effects (i.e., consequences secondary to the ones intended) of the action may be uncertain. For example, it is unclear whether redirecting the trolley will definitely kill the man on the other track. The third aspect concerns the success of the action in preventing the harm. For instance, we would not know for sure whether pushing a man in front of the trolley would successfully stop it.

If we want to gain a better understanding of how people make moral judgments in the real world, moral judgments should thus be considered as an instance of judgment and decision making under uncertainty—not under certainty or only risk. Furthermore, if moral judgments under uncertainty differ from those under certainty, the utility of using moral dilemmas in which all outcomes are certain is unclear. On the one hand, moral judgments could be influenced by people's subjective probabilities about what was foreseeable at a given time if the potential outcomes are weighted accordingly (as in variants of consequentialism). From this perspective, the epistemic situation, and thus the distinction between situations of uncertainty, risk, and certainty, may still be negligible for investigating moral judgments because the decision processes underlying moral judgment could be described in the same way (e.g., Friedman, 1976). On the other hand, research on bounded rationality (Simon, 1955, 1979) has shown that in situations in which information is limited and no objective probabilities are known, people often use simple decision rules (heuristics) which do not weight outcomes by probabilities (Brandstätter, Gigerenzer, & Hertwig, 2006; Dawes, 1979); or even do not rely on probabilities, objective or subjective at all (Gigerenzer & Gaissmaier, 2011). Despite this, uncertainty has received little attention in the domain of moral judgment and decision making (Christensen & Gomila, 2012).

Moral Judgments under Uncertainty and Certainty as Judgments Before and After the Fact

In the present study, we investigate how uncertainty affects moral judgments in dilemma situations, as compared with judgments under certainty. Specifically, we manipulate whether morally undesirable side effects of an action are uncertain or certain by adapting the *hindsight paradigm* (Fischhoff, 1975). In research on hindsight, participants' task is to estimate the probability of a target event, with some people being informed about the actual outcome that occurred, and others not (Hawkins & Hastie, 1990). Two experi-

mental procedures have been employed. The *hypothetical design* compares two groups of participants: one group without outcome knowledge and one group with knowledge of the actual course of events, who are asked to ignore this information. The *memory design* compares participants within one group, contrasting their responses without outcome knowledge to their recall of the original response after outcome information has been provided.

We adopt the hypothetical design which, on the one hand, avoids repeated measure effects, yet on the other hand still allows a realistic implementation of moral judgments in two types of epistemic conditions. In *foresight*, people do not have outcome knowledge and make moral judgments under uncertainty; in *hindsight*, people make judgments after the fact with outcome knowledge, that is, under certainty. Accordingly, in our *foresight* condition (=uncertainty) it is still uncertain whether any negative side effects will occur, whereas in the *hindsight* conditions (=certainty) it is known whether negative side effects occurred (*hindsight bad*) or did not occur (*hindsight good*). Otherwise, the information presented is identical in all conditions, including some uncertainty about the harm the action is supposed to prevent and how successful it may be in achieving this goal.

Using the hindsight paradigm allows a realistic implementation of the epistemic situations of uncertainty and, in particular, certainty. In real world cases, certainty about a course of events only exists after the fact. Thus, the distinction between moral judgments under uncertainty and certainty maps naturally onto the difference between judgments before and after the outcomes of a decision are known. In contrast, the omniscient perspective of trolley-like cases, where all consequences of action and inaction are perfectly certain, does not match the situation of judgments in *foresight* or in *hindsight*—because information about alternative courses of events is rarely available, even retrospectively (Bennis et al., 2010). In addition, the hindsight paradigm has methodological advantages over using dilemmas where everything is certain. For instance, Greene et al. (2009) found lower responsibility ratings in cases where people subjectively expected an action to be less successful than described in the dilemma. Instead of believing the success of the action to be certain as described, some people brought to the task their own beliefs about the likelihood of success, which in turn influenced their judgments (which Green et al. called “un-

conscious realism").[4] In contrast, informing people about the actual outcome that occurred makes it possible to investigate how certainty affects moral judgments without people doubting the certainty of the event.

How do Moral Judgments Differ in Foresight and Hindsight?

There are good reasons not to expect any differences between moral judgments in dilemmas under certainty and uncertainty. First, the decision to act in a moral dilemma results from a morally good intention, despite the potentially negative side effects of the action. Cushman (2008) showed that judgments of permissibility rely principally on an analysis of mental states such as intention, whereas judgments of blame take more account of harmful consequences. If the main focus is on the intention, people may find the action permissible, independent of whether the undesired side effects of the action occur or do not occur. Second, people may rely on moral rules that are based on features of the action other than its consequences, such as "Do no harm" or whether the harm is directly intended or only a foreseen side effect (Aquinas, 13th c. / 1981; Foot, 1967; Mikhail, 2007). Based on such rules, actions could be judged impermissible independent of whether side effects occur or do not occur.[5] In these cases, moral judgments would not differ between *foresight* and *hindsight* contexts (Figure 1, first panel). The difference between certainty and uncertainty would therefore not matter and could be disregarded in the study of moral judgment and the underlying processes.

However, several theoretical approaches suggest a difference in moral judgment under certainty and uncertainty. These accounts make diverging predictions regarding the expected pattern of judgments in *foresight*, *hindsight bad*, and *hindsight good* (Figure 1). According to a first view, people's moral judgments may reflect a consequentialist view of moral judgment. In this case, their moral judgments will be influenced by their beliefs about the probability of negative side effects: under uncertainty, people may simply weigh the potential consequences of action and inaction by subjective probability beliefs—and these beliefs will differ depending on the epistemic situation. Research on hindsight has

[4] As Bennis et al. (2010) point out, the closed-world assumptions of the dilemmas do not hold in real world situations, and people with protected values may be the least likely to accept them.

[5] Not every rule that forbids a certain type of actions will be unaffected by the consequences. If an action is identified by reference to its consequences and not by features of the act alone, a rule may not apply if the consequences are uncertain. For instance, if you do not know for sure whether an action will kill someone, will the rule not to kill apply?

demonstrated across a variety of tasks, domains, and judgments that people adjust likelihood estimates, confidence ratings, or numerical estimates in the direction of already known outcomes, even when they are instructed to ignore them (Christensen-Szalanski & Willham, 1991; Guilbault, Bryant, Brockway, & Posavac, 2004). Theoretical explanations suggest that outcome information may affect the evaluation of original evidence, leading to an automatic update of knowledge used in the judgment (Hoffrage, Hertwig, & Gigerenzer, 2000), or even change the generic mental model used for the domain under consideration (Fischhoff, 1975). From this perspective, hindsight effects represent a by-product of learning mechanisms and are adaptive within the complex and uncertain environments in which people function (Hawkins & Hastie, 1990). If people's subjective probabilities are influenced by knowledge of the outcome, a consequentialist account will predict different moral judgments in foresight and hindsight, because people weigh the potential consequences of action and inaction by their adjusted subjective probabilities to arrive at a moral judgment. The consequentialist account thus leads to three predictions (Figure 1, second panel):

Hypothesis 1a: In foresight, probability estimates of negative side effects are lower than in a condition in which negative side effects occurred (hindsight bad), and higher than in a condition in which negative side effects did not occur (hindsight good).

Hypothesis 1b: Moral judgments are more permissible in foresight than in hindsight bad (where negative side effects occurred), and less permissible than in hindsight good (where no negative side effect occurred).[6]

Hypothesis 1c: The more likely negative side effects are expected to be, the less permissible moral judgments are.

According to a second position, different information may be highlighted in *foresight* and *hindsight*, such that people perceive the situation differently and draw on different cues when making a judgment. For example, Gino, Moore, and Bazerman (2009) used a similar paradigm to investigate *outcome bias* (Baron & Hershey, 1988) in ethical judgments. Outcome bias here refers to moral judgments that take the actual outcomes into account. These are irrelevant to the moral quality of the decisions because they could not have been known at the time of the decision. However, the situations they studied were no

[6] Note that the hypothesis needs to assume that value of consequences is constant across conditions, i.e. that the value of consequences is not itself subject to hindsight bias.

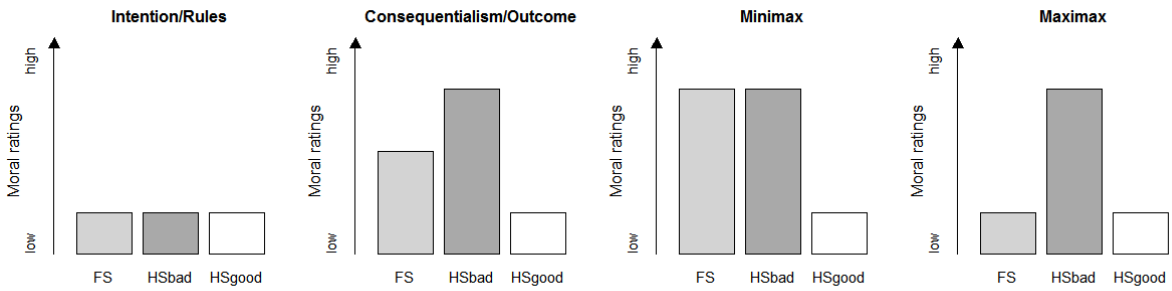


Figure 1. Expected qualitative pattern of moral judgments across *foresight* (FS), *hindsight bad* (HSbad), and *hindsight good* (HSgood) conditions according to different theoretical positions. The consequentialist view makes identical predictions as the outcome bias view as long as probability estimates shift across conditions, as found in hindsight research. If probabilities shift otherwise, the pattern in moral judgment would shift accordingly.

moral dilemmas but cases of potential negligence: Participants rated how unethical it is to not take precautions for the safety of others against rare, but harmful consequences. For instance, in one scenario, a sewage treatment plant has to divert the sewage to a holding tank during a system shut-off. The company decides against investing in expensive back-up systems that would eliminate the risk of overflow—and thus of serious environmental and health consequences—in case of heavy rain. In a within-subjects comparison, people judged this decision to be less ethical when they knew that negative consequences occurred (i.e., in *hindsight*) than without outcome knowledge (i.e., in *foresight*). Conversely, they judged the decision to be more ethical when they knew that no negative consequences occurred than without outcome knowledge. The authors did not attribute the difference to a hindsight effect[7] (i.e., to adjusted probability estimates which in turn influenced the moral evaluation), but to an intuitive decision process in which people judge in *hindsight* based on the actual outcome (see also Gino, 2008)—in contrast to rational deliberation, which would prevent them from taking the outcome into account.

The conclusions drawn by Gino and colleagues (2009) seem too strong in several respects, however. First, in reality, outcomes usually do not occur by chance and may thus provide useful new information. Updating one's beliefs about the likelihood of events in

[7] Gino and colleagues found the same effect in an artificial situation where outcomes were determined by a die. Because the mechanism is known in this case, they argued that people had no reason to adjust their probability estimates and that the effect was thus necessarily due to the outcomes only.

the light of outcome information does not need to be a bias—yet these beliefs were not controlled for. Second, it is neither clear that people rely on rational deliberation in the absence of outcome information, nor do we learn anything about judgments with outcome information by attributing it to an intuitive process beyond a re-description of the observed phenomenon (Kruglanski & Gigerenzer, 2011). If we nevertheless extend Gino and colleagues' findings to judgment in moral dilemmas, the predicted pattern across conditions based on such an outcome effect would be the same as that emerging from hindsight research—though it would not be due to a difference in people's subjective probability estimates (Figure 1, second panel).

Hypothesis 2a: Moral judgments are more permissible in foresight than in hindsight bad (where negative side effects occurred), and less permissible than in hindsight good (where no negative side effect occurred).

Hypothesis 2b: Moral judgments vary independent of people's probability estimates.

A third position takes as starting point research on bounded rationality and the observation that in situations of (Knightian) uncertainty people may sometimes follow simple heuristics that do not rely on probabilities at all (Gigerenzer & Gaissmaier, 2011). For instance, *Minimax* describes a decision strategy that compares the worst possible outcomes and chooses the better one, i.e. the one that minimizes the maximum loss (Savage, 1951; Wald, 1945). Although decision making is different from making a moral judgment in a dilemma, an analogous strategy could be applied in the moral context. In moral dilemmas, both options—not to act as well as to act—come at a moral cost: not acting precludes protecting others from potential harm; acting potentially has negative side effects. Moral dilemmas thus involve at least two morally negative outcomes and their unknown likelihoods which are pitted against each other. If people judge based on a "worst-case" consideration similar to *Minimax* when it is uncertain whether an action will have side effects, they may compare the negative consequences of inaction (which is the inversion of the benefits of the action) with the worst possible consequences of the action (i.e., that negative side effects occur) without considering how likely these are. Across conditions, the following pattern would result (Figure 1, third panel):

Hypothesis 3: Moral judgment ratings in foresight are as low as ratings in a situation in which it is known that side effects occurred (hindsight bad). However, in hindsight good, where no negative side effects occurred, the ratings are higher.

Conversely, people could apply a process similar to the *Maximax* rule (Coombs, Dawes, & Tversky, 1970), which compares the best possible outcomes and chooses the better one. Applied to judgments in moral dilemmas, this could mean that people consider the benefits of the action compared with the best possible case in which no side effects occur (Figure 1, fourth panel):

Hypothesis 4: Moral judgment ratings in foresight are as high as ratings in situations in which it is known that no side effects occurred (hindsight good). However, in hindsight bad, where negative side effects occurred, the ratings are lower.

Thus, depending on the rule people follow and the information considered, ratings in foresight may be similar to those in either hindsight bad or hindsight good without relying on probability estimates.

Experiment

The goal of this experiment was to examine whether moral judgments of an action differ depending on whether its negative side effects are uncertain or known for sure. Because certainty exists only after the fact, we compare moral judgments in *foresight* (where it is uncertain whether a negative side effect will occur) with judgments in *hindsight*, where it is known whether the negative side effect did occur (*hindsight bad*) or did not occur (*hindsight good*). To understand the contribution of subjective probability beliefs to moral judgments before and after the fact, we measure participants' subjective estimates that the negative consequences will occur.

Methods

Participants and design. A total of 730 participants (325 women, age range: 18–75 years, $M = 32.67$, $SD = 12.04$) were recruited through the online platform Amazon MTurk and paid $0.50 for participating in an online study (average duration: 4min 15sec). Only candidates from the United States who had already completed more than 100 online tasks and fulfilled at least 95% of the tasks they accepted were selected for participation.[8] MTurk gives researchers access to a large subject pool that is often more representative of the U.S. population than convenience samples used in lab experiments (Berinsky, Huber, & Lenz, 2012). Indeed, studies suggest that data obtained with MTurk do not differ in quality or reliability from data obtained from traditional lab studies (Buhrmester, Kwang, & Gosling, 2011; Paolacci, Chandler, & Ipeirotis, 2010; Sprouse, 2010). Participants were randomly assigned to one of three between-subjects conditions: *foresight*, *hindsight bad*, or *hindsight good*. To avoid possible order effects in moral judgment (Wiegmann, Okan, & Nagel, 2012), each participant judged only one of six moral dilemmas, resulting in a total of 6 × 3 between-subject conditions.

Materials. We constructed six moral dilemmas based on different real world situations (see Appendix B2 for a full description of all dilemmas). All situations included uncertainty about the threat that an action is supposed to prevent, the desired consequences of the action (i.e., prevention of threat), and potential negative side effects. In order not to omit information that people would typically expect, we asked participants in a pilot test whether they felt any further information was needed in order to make a moral judgment. Based on their responses, we further refined the dilemmas, either by including this information if it would usually be available in a real-life situation or by explicitly stating that it was not available. Each dilemma first introduced the agent (or agency) facing the decision and provided information about the threat. This was followed by a description of the action under consideration, including its potential benefit given the threat as well as its potential negative side effects. This information was identical in all three conditions (*foresight*, *hindsight bad*, and *hindsight good*). In the two hindsight conditions, participants were addi-

[8] MTurk allows researchers to select participants based on their experience with online tasks (i.e., the number of tasks they have already participated in) and their work reliability (i.e., how many tasks they have accepted and in fact completed as required by the researcher).

tionally informed that the action was eventually taken and that the negative side effects did occur (*hindsight bad*) or did not occur (*hindsight good*).

Figure 2 illustrates the structure of the moral dilemma for the three conditions, using a dilemma regarding the use of genetically modified corn (GM corn).

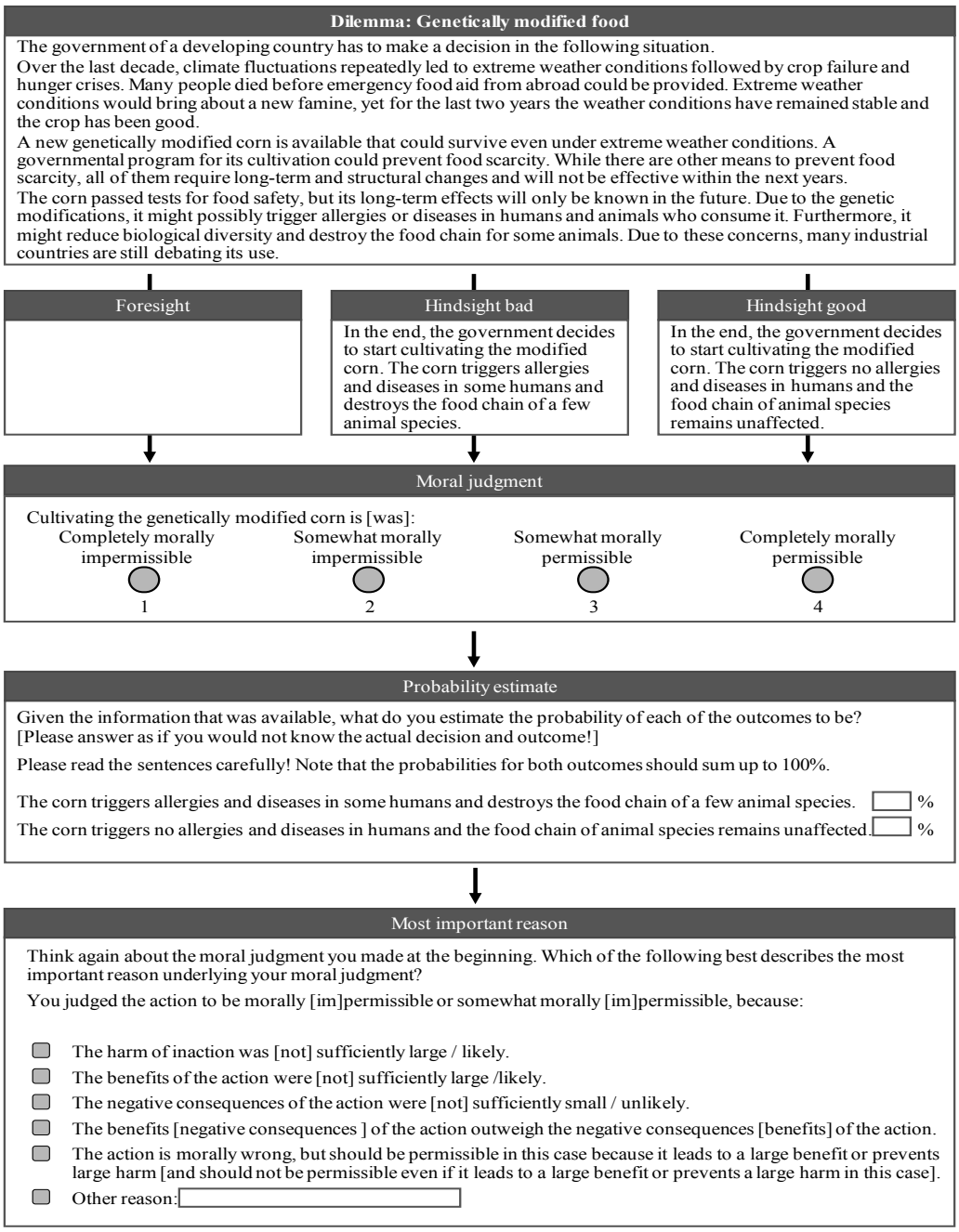


Figure 2. Overview of the experimental design and procedure for the GM corn dilemma. In foresight, it was uncertain whether the negative side effects of the action would occur. In the hindsight conditions, we provided information on whether the action resulted in negative side effects (*hindsight bad*) or not (*hindsight good*).

The top box shows the information that was identical in all conditions, including the description of the threat (imminent famine), the decision under consideration (cultivation of genetically modified corn to prevent future famines), and the possible negative side effects (allergies, destruction of biological diversity and food chains).

In the two hindsight conditions, participants were additionally told that the action was eventually taken and were provided with information on the occurrence of the negative side effects. For instance, in the *hindsight bad* condition, they were told that the government decided to cultivate the genetically modified corn, which later caused allergies and diseases in humans. By contrast, in the *hindsight good* condition, they were told that the negative side effect did not occur, that is, that the genetically modified corn did not cause any allergies or diseases. The manipulation of the other five dilemmas followed the same rationale. The six dilemmas related to influenza vaccination, breast cancer screening, torture, genetically modified corn, dam building, and provision of state credit (see Appendix B2). In each dilemma, the action has the potential to avert a threat, but also potentially entails some negative side effects (Table 1). Participants were instructed to judge the action of each person or agency independently of what current law specifies.

Table 1

Overview of Morally Relevant Aspects in Each Dilemma Situation

Dilemma	Threat	Action	Potential negative side effects
Influenza	pandemia	large-scale vaccination campaign	severe or even fatal side effects of new vaccination
Torture	rescue kidnapped boy in time	threatening to torture to get information	acquittal of kidnapper due to violation of procedural rules
Screening	breast cancer	regular screening	wrong diagnosis and unnecessary treatment
GM corn	famine	cultivating GM corn	allergies and diseases and destruction of food chain of some animals
Dam	severe floods	building a dam	lower water quality and higher costs of drinking water
Credit	loss of 5000 jobs through bankruptcy	reallocate state funds to provide state credit	shortfall of state loans for start-ups

After receiving this information, participants were asked to judge the moral permissibility of the action (see Figure 2). Judgments were given on a scale ranging from 1 to 4 (1 = completely morally impermissible, 2 = somewhat morally impermissible, 3 = somewhat morally permissible, 4 = completely morally permissible). Subsequently, their estimates regarding the probability of the negative side effects were elicited (Figure 2). The goal was investigate whether knowledge of the occurrence of the side effects would influence participants' probability judgments relative to the *foresight* condition in which this knowledge was not available. Probability estimates were always elicited after the moral judgment because asking for the estimates first might have impacted the judgment process. Finally, participants were asked to provide the most important reason for their moral judgment, using a forced choice paradigm (see Appendix B3). The goal was to shed light into the cues participants use when making moral judgments.

Procedure. The experiment was conducted online. On the first screen, participants received a description of the dilemma (see Figure 2, top box). In the two hindsight conditions, this screen also provided information on the occurrence or non-occurrence of negative side effects (*hindsight bad* vs. *hindsight good* condition). After reading the dilemma, participants made their moral judgment on a second page, on which the question and scale appeared below the dilemma. On screen three, they were asked to judge the probability of negative side effects of the action. In the two hindsight conditions, they were instructed to give their estimates as if they did not know the actual decision and whether or not side effects occurred.

The final dependent variable aimed at exploring the cues that participants use in moral judgment. Participants were asked to identify the most important reason for judging the action to be permissible (rating 3 or 4) or impermissible (rating 1 or 2; see Appendix B3). They were given a choice between the harm of inaction, the benefits of the action, its negative consequences, the trade-off between costs and benefits, the reasoning that the action is morally wrong but should be permitted in this case (for participants who judged the action to be permissible) or should not be permitted irrespective of its benefits (for participants who judged the action to be impermissible), and an open-ended option in case none of the reasons applied. The last item was taken from Ritov and Baron (1999) and indicates whether a person holds a protected value—that is, a value not to be traded-off, no

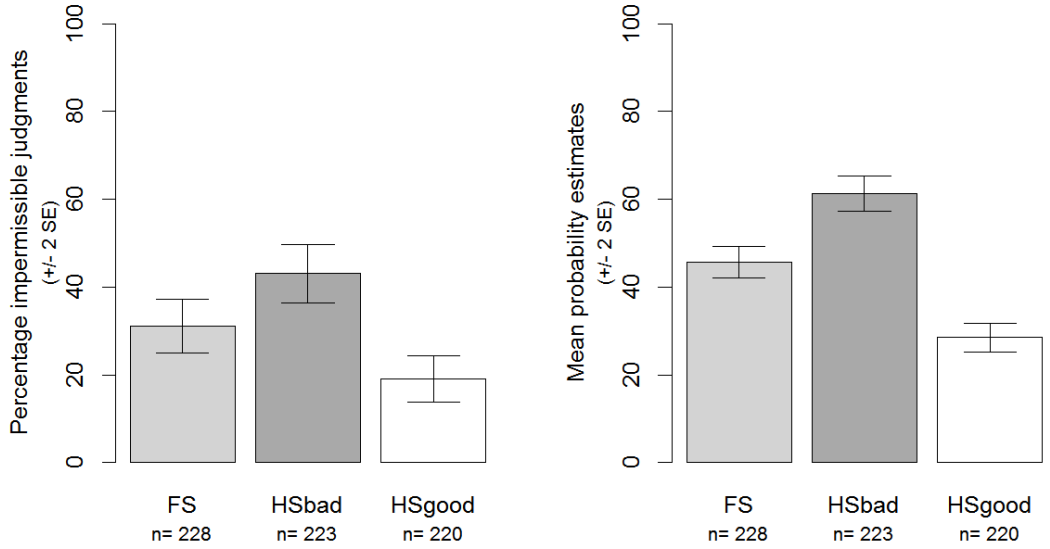
matter how great the benefits are. Finally, each participant was presented with a simple logical task ("If an action A is better than an action B, and B is better than C, which action is the worst?") to exclude those who did not pay enough attention, and asked to provide some demographic information. In addition, all participants were asked whether they had previously participated in the study, while being informed that this would not affect their payment.

Results

Of the initial 730 participants (325 women, age range: 18–75 years, $M = 32.67$, $SD = 12.04$), 14 participants were excluded prior to data analysis because they already participated; another 45 participants were excluded because they failed the attention-verification test. This left 671 participants for the data analysis (297 women, age range: 18–75 years, $M = 32.75$, $SD = 12.18$), with $n = 34$–42 in each of the 6×3 conditions.

Moral Judgments in Foresight and Hindsight

Figure 3 (left panel) shows the percentage of participants in each condition who judged the action to be completely or somewhat morally impermissible, aggregated across all six dilemmas (see Table B1 in the Appendix, for descriptive statistics of individual dilemmas). The results show that participants' moral judgments differed systematically across the three conditions. In the *foresight* condition, 31% of participants judged the action to be somewhat or completely morally impermissible. In the *hindsight bad* condition, in which subjects were told that side effects did occur as a consequence of the action, a higher proportion of participants (43%) judged the action to be morally impermissible. In the *hindsight good* condition, in which side effects did not occur, only 19% judged the action to be impermissible. Thus, in *hindsight* different moral judgments were given than in *foresight*.



F*igure 3.* Proportion of participants rating an action to be completely morally impermissible or somewhat morally impermissible (scale rating 1 or 2) (left panel), as well as mean probability estimates for negative side effects (right panel) in *foresight*, *hindsight bad*, and *hindsight good* conditions aggregated across all dilemmas.

We used non-parametric U-tests to analyze these results.[9] The *foresight* condition differed both from the *hindsight bad* condition ($U = 23055$, $z = -1.83$, $p = .07$, $r = 0.09$) and the *hindsight good* condition ($U = 19124$, $z = -4.7$, $p = .001$, $r = 0.22$). The largest difference was obtained between the two hindsight conditions ($U = 16875$, $z = -6.03$, $p = .001$, $r = 0.29$) (see also Figure B1 in the Appendix). On the level of the individual dilemmas, the same qualitative trend can be seen in five out of six dilemmas (Figure 4), although not all pairwise comparisons were significant in each dilemma (see Table B2 in the Appendix). In the sixth dilemma ("credit"), the action was not judged to be more impermissible in *hindsight bad* than in *foresight*; rather, the percentage of participants rating the action to be impermissible was similar in both hindsight conditions. It is possible that in this case the actual amount of damage in *hindsight bad* (i.e., the number of start-ups declaring insolvency) sounded less disastrous than might have been feared in *foresight*, where the extent of the negative side effect was fully uncertain.

[9] The scale includes a change in category from item 2 (somewhat morally impermissible) to item 3 (somewhat morally permissible). As it is not an interval scale, *t*-tests were not used. A chi-square analysis based on the binarized measure (impermissible vs. permissible) as reported above leads to the same results (see also Figure 3, left panel).

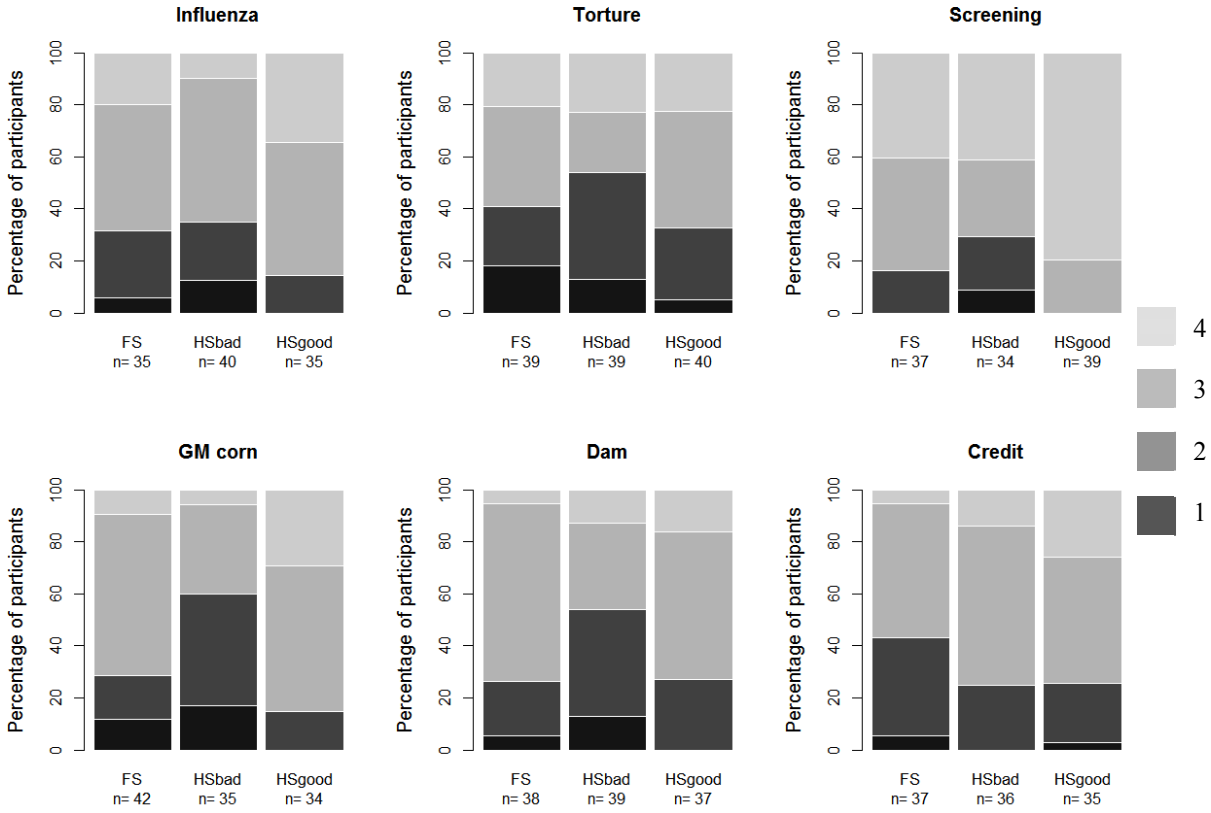


Figure 4. Percentage of participants on each level of the moral judgment scale (1- 4) in *foresight* (FS), *hindsight bad* (HSbad), and *hindsight good* (HSgood) by individual dilemma. In all dilemmas except "credit," the same trend is visible: The proportion of people rating the action to be completely or somewhat impermissible (scale rating 1 or 2) is highest in HSbad, lowest in HSgood, and in between in FS.

Overall, the pattern of moral judgments is consistent with the predictions of a consequentialist account based on adjusted probability estimates, or with an outcome bias view as discussed by Gino et al. (2009). This holds both aggregated across dilemmas as well as on the level of the individual dilemmas.

Probability Estimates in Foresight and Hindsight

Next, we analyzed participants' probability estimates regarding the likelihood of the actions having negative side effects. Research on hindsight shows that people adjust their subjective probabilities in the direction of known outcomes, even if they are explicitly instructed to ignore this information (Fischhoff, 1975). Figure 3 (right panel) shows participants' probability judgments for the three conditions, averaged across all dilemmas (see Table B3 in the Appendix, for descriptive statistics of individual dilemmas). Interestingly, the observed pattern corresponds to that found for moral judgments, with participants judging side effects to be most likely in the *hindsight bad* condition (M = 61.4, SD = 30.2), least likely in the *hindsight good* condition (M = 28.5; SD = 23.5), and the *foresight* condition being in between these estimates (M = 45.7, SD = 27.3). All differences were statistically reliable: *foresight* vs. *hindsight bad* (U = 17,747, z = - 5.56, p = .001, r = -0.26), *foresight* vs. *hindsight good* (U = 15,937, z = - 6.69, p = .001, r = -0.32), as well as the two hindsight conditions (U = 9,967, z = -10.84, p = .001, r = -0.51).[10] The same pattern was observed for each individual dilemma (see Table B4 in the Appendix B,), again with the exception of "credit," in which probability estimates in the *foresight* condition were rather high and similar to those in the *hindsight bad* condition (Figure 5).

That negative side effects were rather more expected in the "credit" scenario again indicates that the diverging pattern of moral judgments in this case may be due to uncertainty about the extent of negative side effects. Overall, knowledge of the actual course of events was mirrored in participants' probability judgments, even though they were explicitly instructed in the hindsight conditions to ignore the decision and resulting outcome (see Figure 2). Taken together, these findings indicate a hindsight effect for probability estimates in the context of moral dilemmas.

[10] Given what we know from research on risky choice, people's probability estimates are not well described by an interval scale, so we again used a Mann–Whitney test instead. However, a t-test on the means gives the same results.

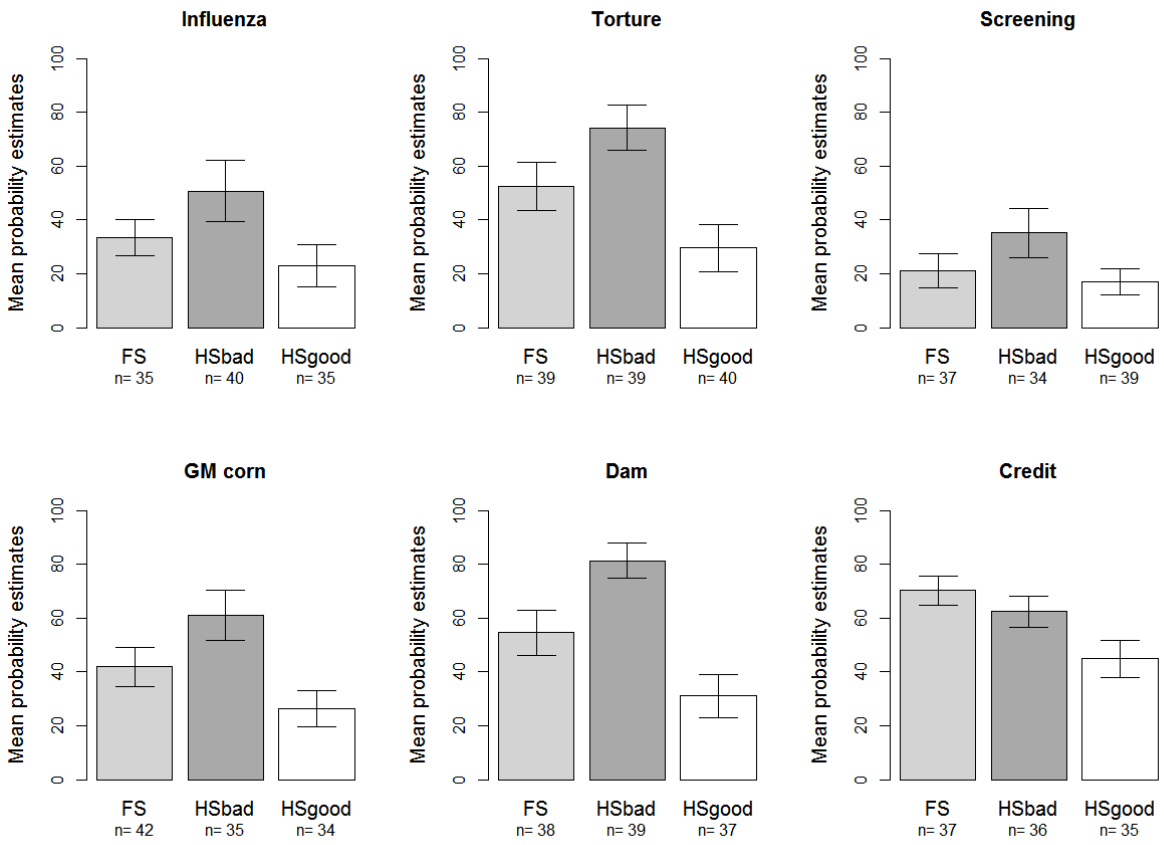


Figure 5. Mean probability estimates that negative side effects will occur in the *foresight*, *hindsight bad*, and *hindsight good* conditions by individual dilemma. Error bars indicate 2 standard errors above and below the mean.

Relation Between Moral Judgments and Probability Estimates

Is there a systematic relation between participants' moral judgments and their probability estimates of negative side effects? If so, actions should be judged less permissible the more likely the negative side effects are considered to be. Figure 6 plots the mean probability estimates as a function of moral judgment, aggregated across dilemmas.

Within each condition, a systematic trend is observed: participants who considered the action to be completely impermissible gave in the aggregate higher probability estimates, with estimates decreasing as a function of moral permissibility, and people rating the action to be completely permissible giving the lowest probability estimates (*foresight*: $H(3) = 23.74$, $p = .001$; *hindsight bad:* $H(3) = 27.72$, $p = .001$; *hindsight good:* $H(3) = 18.21$, $p = .001$). In line with this, there is a moderate negative mean correlation between moral judgment and probability estimates (*foresight* $r_s = -.32$; *hindsight bad* $r_s = -.35$; *hindsight good* $r_s = -.29$). Thus, judgments and probability estimates correspond to each

other as predicted by a consequentialist account of moral judgment, according to which people weigh potential outcomes by subjective probabilities adjusted in the direction of the actual outcome. (Note however that it is also possible that people adjusted their probability estimates to fit their moral judgments.)

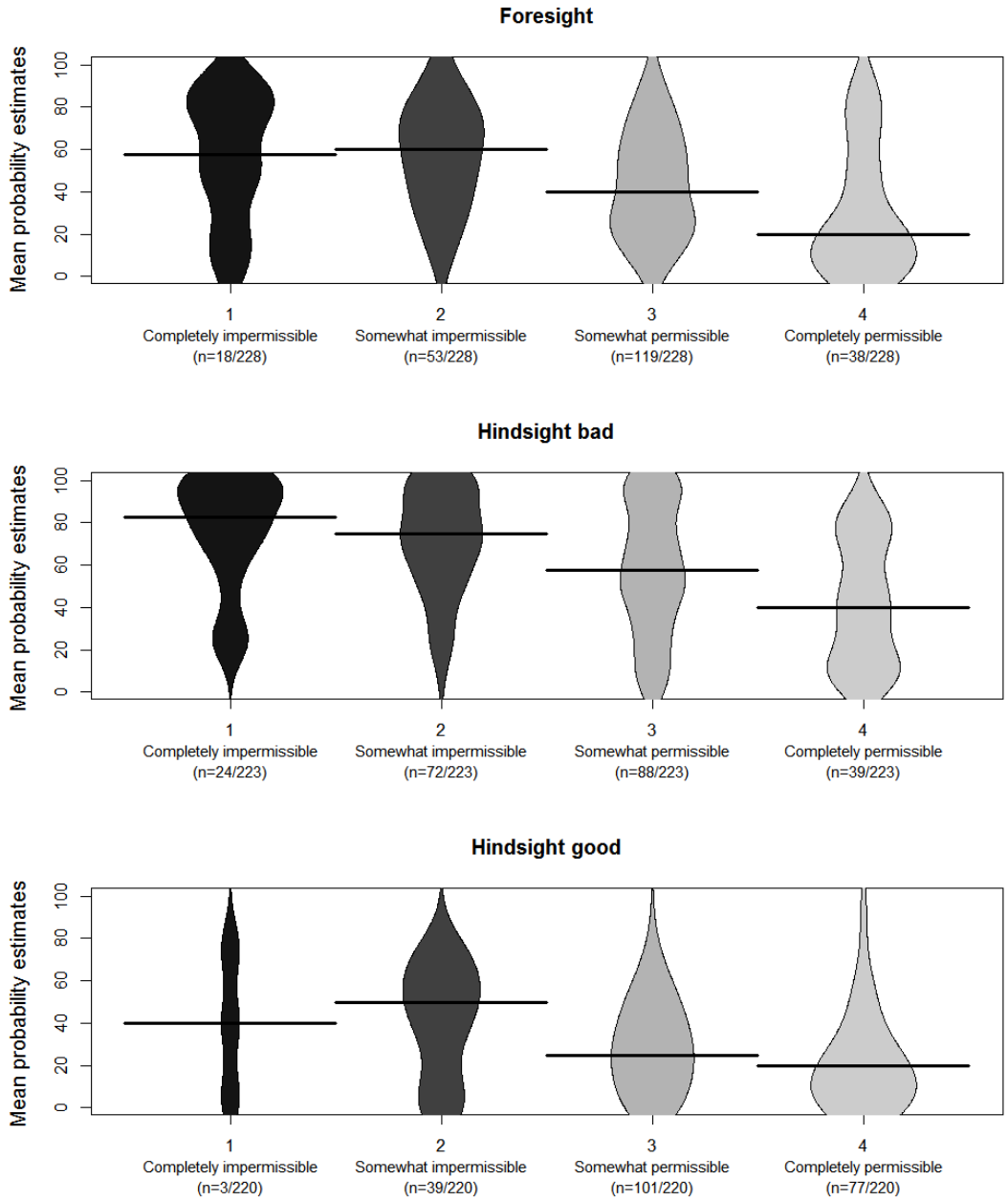


Figure 6. Beanplots of probability estimates by moral judgment rating in *foresight*, *hindsight bad*, and *hindsight good* conditions aggregated across all dilemmas. The shape of the beans displays the density of the distribution, with the horizontal lines indicating the median.

However, the correlation between both measures is only moderate, which suggests that not everyone takes a consequentialist approach. Inspection of individual dilemmas (see Table B5 in the Appendix) reveals that correlations vary across conditions, indicating where probability estimates differed while moral ratings did not. Finally, the overall correlations for individual dilemmas across conditions varied from weak (torture) to strong (GM corn), showing that the relevance of subjective probabilities for moral judgments—and thus of the underlying judgment process s—differs depending on the situation (see Discussion).

Reason Analysis

To further explore the cues people use and as a starting point for process models of moral judgments, we asked participants to identify the most important reason for their judgment by selecting one of five reasons or an open ended option (see Figure 2). The available reasons reflect abstract categories of cues that can be found in each dilemma independent of the particular context: the harm of inaction, the benefits of the action, its negative consequences, the trade-off between costs and benefits. The fifth item stated that the action is morally wrong and is only permissible in this exceptional case (for participants who judged the action to be permissible), or is not permissible in any case (for participants who judged the action to be impermissible). Participants selecting this item were identified as holding protected values (Ritov & Baron, 1999).

Does the distribution of reasons allow for some exploratory insights into which cues participants focused on? Figure 7 shows a normalized histogram of the reasons participants selected as most important for their moral judgment in *foresight*, *hindsight bad*, and *hindsight good*, aggregated across all dilemmas. In *foresight*, the most often selected reason was a cost–benefit trade-off with respect to the action. This choice may reflect the fact the structure of a moral dilemma highlights both the costs and benefits of an action. The next most frequently stated reasons were the benefits of the action and the threat it may prevent. Both reasons could be taken as two sides of the same coin: whereas looking at the benefits shows a focus on what can be gained by acting, a focus on the threat shows a concern about a need for prevention. The least selected options were the cost of the action and moral reasons. Whereas people who selected the cost category focused on the negative side effects, those choosing the moral category perceived a moral cost beyond these negative

consequences. Overall, in *foresight* more people stated that their judgment was primarily driven by a concern with the need for or benefits of prevention than by a concern with costs only.

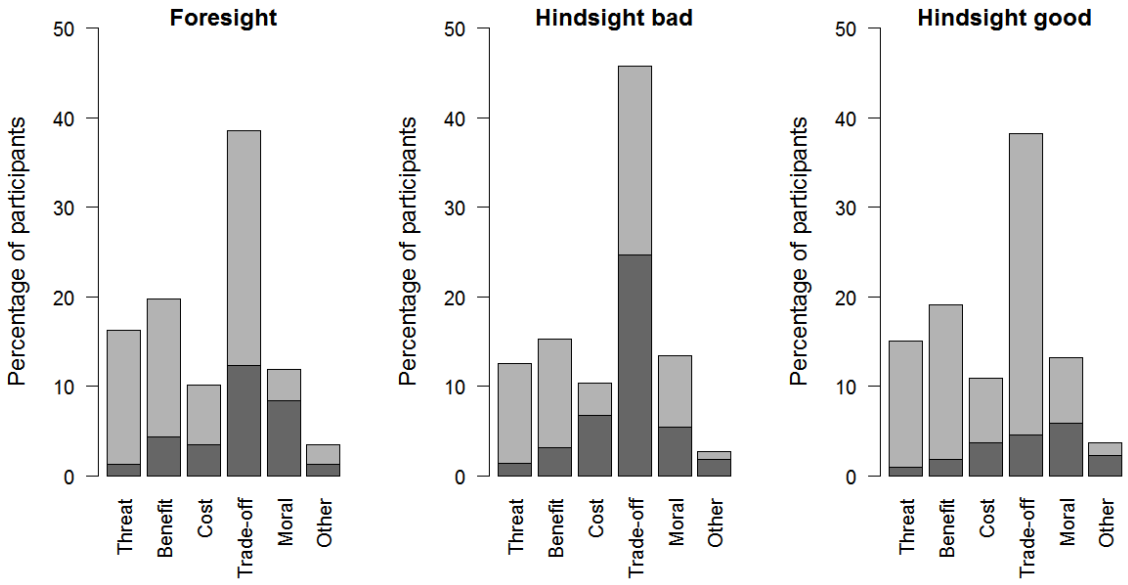


Figure 7. Normalized histogram showing the proportion of participants who selected each reason as being most important for their moral judgment by *foresight*, *hindsight bad*, and *hindsight good* condition separately. The colors indicate whether people judged the action to be somewhat or completely permissible (light-grey) or somewhat or completely impermissible (dark-grey) based on that reason.

While basically the same picture emerged in *hindsight good*,[11] the distribution of reasons observed in *hindsight bad* differed. Here, the overall endorsement of the trade-off category was decisive for almost half of the participants. Furthermore, the proportion of participants for whom the trade-off provided a reason to judge the action impermissible doubled from 12% in *foresight* to 25% in *hindsight bad*. At the same time, endorsement of the benefits and threat categories was less frequent than in *foresight*, so that the concern with a need for or benefits of prevention was no longer more important than a concern with costs only. The different distribution of reasons in *hindsight bad* shows that side effects are

[11] The main difference between f*oresight* and *hindsight good* was whether participants judged an action to be permissible or impermissible based on the reason they selected as most important. Because the total number of impermissible judgments is lower in *hindsight good* than in *foresight*, the proportion of people who judged the action impermissible based on each reason is of course lower as well.

evaluated differently when it is known that they occurred, which makes other cues, such as the threat and the benefits of the action, loose in importance.

How do reasons interact with moral hindsight? If the reasons people state really indicate whether they took potential or actual side effects into account in their judgments, differences in moral judgments across *foresight* and *hindsight* should only be found conditional on the reasons they stated. Specifically, participants who stated that they based their judgment primarily on the cost–benefit trade-off or the costs of the action should be sensitive to the likelihood of negative side effects in their moral judgments (n = 344). In contrast, people who were concerned with the need for or benefits of prevention or who considered moral costs beyond the side effects should be insensitive to our manipulation (n = 305). The remaining 22 participants, who selected other reasons, were excluded from the following analysis.

Aggregated across all dilemmas, we found the same hindsight effect in probability estimates that was found for the total sample in both subgroups (see Figure 8). However, the hindsight effect in moral judgments is absent for the *insensitive* subgroup. Consequently, the hindsight effect is even larger in the *sensitive* subgroup than in the total sample. Thus, whereas the hindsight effect on probability estimates is found within each subsample, the effect on moral judgments only emerges for people who used reasons sensitive to the negative consequences of the action. Thus, independent of the moral judgment people made, the reasons they stated as most important for their judgment actually bring insights into the cues that informed their judgment strategy and provide a starting point for the development of process models for moral judgment.

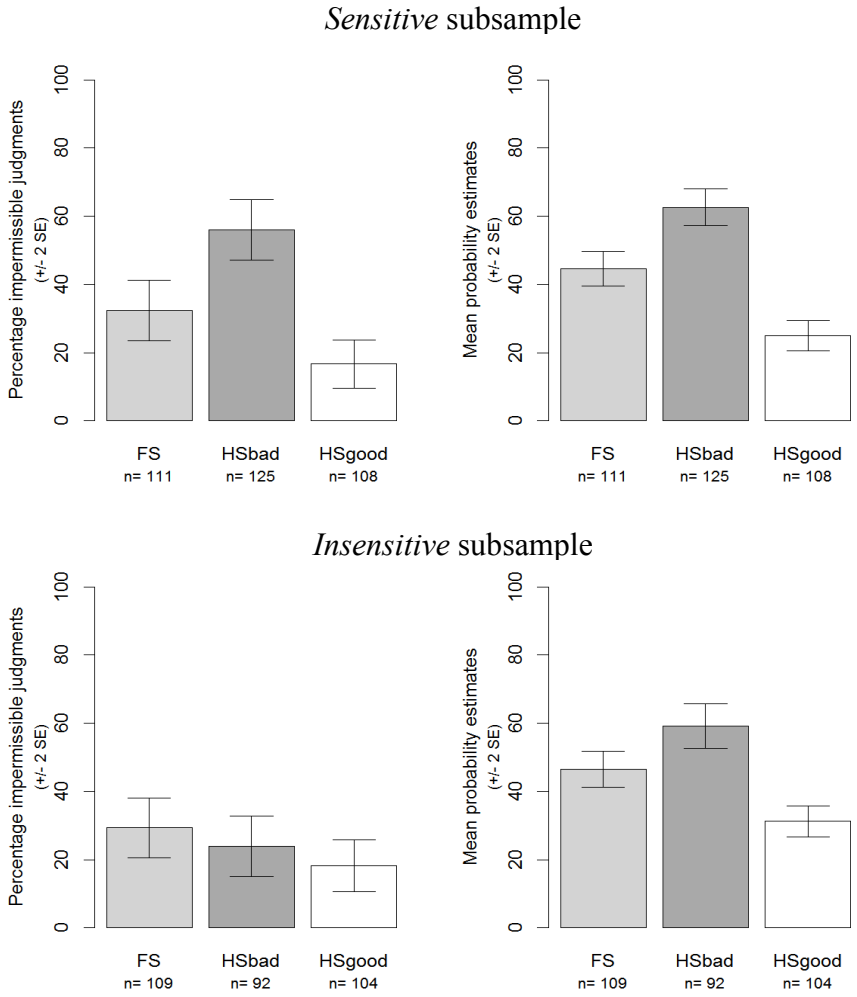


Figure 8. Comparison of the subsample of participants who stated reasons sensitive to negative consequences (upper panel) vs. insensitive to negative consequences (lower panel). The left panel shows the proportion of participants rating an action to be completely morally impermissible or somewhat morally impermissible (scale rating 1 or 2), the right panels shows the mean probability estimates for negative side effects. The hindsight effect in probabilities was found within each subsample, yet the hindsight effect for moral judgments emerged only for the sensitive subsample. Error bars indicate 2 standard errors above and below the mean.

Discussion

Moral judgments are often made under some uncertainty about the situation, as certainty only exists after the fact. The goal of the present study was to compare moral judgments in two kinds of epistemic situations: when it is uncertain whether an action will lead to negative side effects (*foresight*), and when it is already certain that negative side effects

did occur (*hindsight bad*) or did not occur (*hindsight good*). Our key finding is a hindsight effect in moral judgments. Participants in the *foresight* condition judged actions to be more permissible than did participants in the *hindsight bad* condition, where negative side effects occurred. Conversely, they judged actions to be less permissible than did participants in the *hindsight good* condition, where negative side effects did not occur.

The second finding is a hindsight effect in participants' probability estimates within the context of a moral dilemma. Consistent with research on hindsight, side effects were estimated to be more or less probable depending on whether or not they occurred. This suggests that the hindsight effect in moral judgments is different from an outcome bias (Gino et al., 2009)—that is, from a direct effect of the actual outcomes, independent of a change in subjective probabilities. Third, judgments and estimates correspond to each other as predicted by a consequentialist approach, which weighs potential outcomes by subjective probabilities adjusted in the direction of known outcomes. This interpretation is further supported by the exploratory analysis of participants' reasons: 366 out of 671 participants gave reasons that take side effects into account, with a cost–benefit trade-off being the most often stated reason (approx. 40%) in each condition.

However, there are three points of qualification. First, it is important to note that the hindsight effect in moral judgments was weaker than that for probability estimates, and that the correlation between the two measures was only moderate. This suggests that the perceived likelihood of side effects mattered only for some participants' moral judgments, whereas others followed a process that does not rely on these cues. Data on the most important reasons for choices provided a helpful way to classify people, independent of their moral judgments, into those who attended to cues relating to side effects and those who did not. In addition, these analyses provided some insights into the proportion of people for whom different cues were important. The exploratory analysis indicated that across all three conditions about half the participants (305 out of 671) endorsed reasons that were unrelated to the side effects of the action. For this subgroup, no hindsight effect in moral judgments was found across conditions, although there was still a strong effect for probability estimates. Thus, a hindsight effect in moral judgments does not generally result from a change in outcome information, but from the interaction with processes that rely on this information.

The finding that not everyone follows a consequentialist policy is consistent with previous research on moral dilemmas: depending on the situation, more or less people will instead follow some kind of deontological rule or further features of the action other than its consequences, such as whether harmful consequences are directly intended or only a foreseen but unintended side effect (Aquinas, 13th c. / 1981; Foot, 1967). For instance, in the trolley dilemma, most people do not find it permissible to save five people on the tracks by throwing another person in front of a trolley, although many do find it permissible to redirect the trolley even if it will kill another person. In contrast to a consequentialist process, which takes outcomes as well as likelihoods into account, deontological rules or processes, which focus on features of the action, can require less information and may be more stable under varying epistemic conditions—and thus also more robust against hindsight effects. It would therefore be interesting to investigate how far the deontological rules and other important features of actions that research on moral dilemmas had identified interact with realistic epistemic conditions in *foresight* and *hindsight*. However, to better understand how moral judgments under uncertainty differ from moral judgments when outcomes are known, researchers would need models of the decision process that more precisely specify which cues are processed, and how.

Second, while the hindsight paradigm is one way to implement judgments under certainty and uncertainty in an ecologically valid way as compared with dilemmas in which all outcomes are certain, the method also has its limitations. For example, within the paradigm it is impossible to hold probability estimates constant across *foresight* and *hindsight*, so it cannot be determined whether a change in probability estimates is really a necessary condition for a change in moral judgment. One solution would be to provide the same probability information explicitly in *foresight* and *hindsight* while changing outcome information—but this would no longer be a situation of uncertainty, but one of risk (Knight, 1921). And while some may argue that there is no relevant difference as people rely on subjective probabilities (Friedman, 1976), this claim is rather part of the question at stake and should thus be empirically studied rather than being presupposed.

Third, knowing the actual outcome after the fact is of course not the same as being certain about which events will occur *in foresight*. However, in reality, full certainty only occurs after the fact, and cases that come close to certainty in *foresight* are just one end of the full range of epistemic conditions that people encounter. Thus, if we want to under-

stand how people make moral judgments, it may be more worthwhile to study epistemic conditions that are representative—rather than to use artificial dilemmas in which all outcomes are certain.

Fourth, it would be important to determine whether the observed hindsight effect generalizes to other measures of moral judgment, such as judgments of blame, punishment, or responsibility. For instance, Walster (1967) found a hindsight effect in people's confidence of having foreseen a financial loss—yet no corresponding effect on the attribution of responsibility. In contrast, Cushman (2008) suggests that judgments of blame rely more on consequences and causation than do judgments of permissibility. Thus, hindsight effects in, for example, evaluations of negligence should loom even larger than in the present study.

We suggest that it is important to include uncertainty and other epistemic conditions that people face in the real world when investigating the processes that they rely on in moral judgment. As moral judgments differed in *hindsight* and in *foresight*—and thus under certainty and uncertainty—the common use of dilemmas in which everything is certain is not an unproblematic simplification in the empirical study of moral judgments. Taking different epistemic conditions into account is also important from a normative perspective, however. Although it may often be adaptive to adjust one's beliefs based on outcome information, it seems inappropriate to judge a decision maker based on knowledge about the actual outcomes if the evidence at the time of the decision did not allow it to be foreseen (Zimmerman, 2010). Furthermore, a better understanding of how people arrive at their moral judgments may be useful for normative theories, to provide practical guidance not only in principle, but in practice—that is, under realistic epistemic conditions. This study is just a first step towards a better understanding of the interaction of moral judgments and the epistemic conditions under which they are made. An open question is whether the same effect will be observed for other potentially uncertain aspects of the dilemma. For instance, the diverging pattern of results for the "credit" dilemma indicates that moral judgments may also reflect uncertainty about the size of outcomes, such as the magnitude of negative consequences, rather than uncertainty about whether or not they will occur. In the end, the dilemmas used in this study still represent "small worlds," in that options and consequences were given. In real cases, the search for—and the awareness of—possible consequences and alternative options may often present the greatest challenge for the improvement of moral judgment and behavior.

Chapter 5

General Discussion

Uncertainty pervades many situations in which moral judgments and decisions are made. Yet research on moral judgment and decision making typically has paid little attention to the epistemic conditions we face—and thus to the various degrees of uncertainty under which we have to judge and decide.

In contrast, the research projects of this dissertation investigated how people make judgments and decisions in moral situations under uncertainty. By adopting the perspective of bounded and ecological rationality, the work presented here emphasizes the need to study judgment and decision making under epistemic conditions representative of those encountered in real life. In this last chapter, I will briefly summarize the insights gained in this dissertation and point to open challenges for future research.

What Have We Learned?

The dissertation presented theoretical and empirical work seeking to understand judgment and decision making in moral situations. Chapter 2 adopted a broader theoretical approach and analyzed judgment and decision making in moral situations from the perspective of ecological rationality (Todd et al., 2012). Chapter 2 and 3 reported empirical studies that focused on risk and uncertainty in two research paradigms frequently used in

the study of moral behavior and moral judgment, respectively. While Chapter 3 investigated cooperation in social dilemmas under different levels of risk and with information about risk being acquired in different ways, Chapter 4 compared judgments in moral dilemmas under certainty and uncertainty. What are the insights we have gained and the lessons we have learned from studying moral judgment and decision making from this perspective?

The same heuristic may explain systematically different outcomes, ethical or unethical, depending on differences in the environment. Chapter 2 discussed how the perspective of ecological rationality can provide a parsimonious explanation of three types of moral inconsistencies that often have been observed in moral judgments and decision making. In contrast to approaches that explain moral judgment and behavior by characteristics of the individual, such as character traits (Doris, 2002), social preferences (Fehr & Schmidt, 1999), or moral stages (Kohlberg, 1984), the interaction of heuristics and the environment can help to explain systematically different outcomes, ethical and unethical, depending on differences in the environment. Instead of resorting to ad hoc explanations or mere data fitting, the approach allows not only for a more parsimonious explanation but for actual predictions of behavior that appears inconsistent from an internalistic point of view.

The perspective of ecological rationality further leads to a functionalist understanding of moral judgment and decision making as a means to regulate social relations (Haidt & Kesebir, 2010; Rai & Fiske, 2011). In line with this, we argued that people may often rely on social heuristics that serve the social coherence of groups rather than on particular moral norms or rules. Such social heuristics may allow them to deal efficiently with epistemic uncertainty, while at the same time serve important social goals. Importantly, from an ecological rationality perspective, a heuristic may serve these functions, yet not necessarily be ethically appropriate according to a normative moral theory or the ethical goals we adopt.

Understanding judgment and behavior in moral situations requires investigation of the uncertain and social environments in which they are made. As a methodological consequence, the perspective of ecological rationality implies that we need to take the epistemic as well as the social environment into account. If we want to understand judgment and behavior in moral situations, we need to study environments that not only

show the epistemic and social uncertainty common to the real world, but that also allow people to rely on social heuristics under uncertainty (see, e.g., Wit, van Dijk, Wilke, & Groenenboom, 2004).

Decision making in social dilemmas depends on differences in risky environments but not on the way risk information is acquired. Given the importance of epistemic and social uncertainty, Chapter 3 empirically investigated cooperation decisions in social dilemmas that explicitly included both sources of uncertainty. When cooperating, people face social uncertainty about the action of others because the benefits of cooperation depend on the behavior of others. Yet even if everyone cooperates, the results of such joint efforts are often uncertain due to risky environments.

While epistemic uncertainty that stems from the environment (often referred to as "environmental uncertainty") has to some extent been included in the study of social dilemmas (for a review, see von Dijk, Wit, Wilke, & Budescu, 2004), different levels of risk and the way information about it is acquired have not previously been investigated. Typically, risk information has been provided by a description of probabilities—without taking into account that experiencing risk not only comes closer to decisions made in real life contexts, but may also trigger different decision processes (Hertwig, Barron, Weber, & Erev, 2004). We compared cooperation decisions in social dilemmas based on a description of environmental risk with decisions based on experience. Furthermore, we investigated whether cooperation is sensitive to differences in risky environments by varying the combination of outcomes and probabilities for the same expected outcomes. Results show that cooperation varied systematically across risky environments. Instead of resulting from social preferences (e.g., Fehr & Schmidt, 1999) or other characteristics of the individual, cooperation systematically depended on differences in risky environments. However, the way information was acquired did not make any difference for cooperation, but only for nonsocial lottery decisions. Thus, while the level of environmental risk was identical in social dilemmas and in lotteries, a DE gap was found only in the latter. This discrepancy between the social and the nonsocial decision situation again underlines the importance of studying moral decision making under conditions that match the epistemic and social environments people encounter outside the lab.

Moral hindsight: moral judgments differ under certainty and uncertainty. Chapter 4 focused not on decision making but on judgments about the decisions of others in moral dilemmas. In contrast to typically investigated moral dilemmas that present all future consequences as certain (e.g., Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Hauser, Cushman, Young, Kang-Xing Jin, & Mikhail, 2007; Mikhail, 2007), we compared judgments in foresight, when the side effects of an action were still uncertain, to those in hindsight, when it was known whether side effects occurred. The key finding was a hindsight effect (Fischhoff, 1975) in moral judgments. That moral judgment differed under certainty and uncertainty shows the importance of taking uncertainty into account when examining the processes underlying moral judgments. Furthermore, it raises the question of the extent to which results gained by using dilemmas in which everything is presented as certain will generalize to judgments under realistic epistemic conditions.

Second, we also found a hindsight effect in the probability estimates people gave for negative side effects. Both results together are consistent with a kind of consequentialist judgment process that weights the potential consequences by probabilities adjusted in the direction of already known outcomes. Nevertheless, the correlation between moral judgments and probability estimates was only moderate. Together with an exploratory analysis of the reasons people indicated as important for their judgments, these results suggest that only some people followed a consequentialist process, while others relied on other cues, such as the benefits of the action or the threat that needed to be prevented.

That people rely on different decision strategies further underlines the importance of considering the epistemic conditions. On the one hand, these may influence the decision processes that people rely on or even those that are applicable under uncertainty. On the other hand, simple heuristics have been shown to be more robust under limited information than "rational" models (Gigerenzer & Brighton, 2009). Similarly, some heuristics in the moral domain may be less sensitive to uncertainty and limited information than a consequentialist decision process and thus more robust against hindsight effects.

What Do We Still Need to Learn?

The work presented in this dissertation contributes to the study of morality by adopting the perspective of bounded and ecological rationality. This approach emphasizes the interplay of cognitive processes and the structure of the environment—and thus the uncertainty under which moral judgments and decisions have to be made. While the studies provided valuable insights, the approach calls for further empirical as well as theoretical work to develop a better understanding of the decision and judgment processes employed in moral situations. Instead of re-summarizing the more detailed questions for future research that arise from each project, I will conclude by pointing out three fundamental challenges that have yet to be met if we are to gain a better understanding of judgment and decision making in moral situations.

Beyond small worlds: How to make moral judgments and decisions when information about states of the world, consequences, or alternative options is missing? The theoretical and empirical results of this dissertation show the importance of considering different sources and degrees of uncertainty if we want to understand judgments and decisions in moral situations. However, the research presented here is only a small step towards a better understanding of the interaction of moral judgments and the epistemic conditions under which they are made. In real cases, epistemic uncertainty results not only from the fact that the course of events is uncertain. Often, we will not even be aware of possible states of the world or the alternative options available (for a taxonomy of different sources of uncertainty, see Brighton & Gigerenzer, 2012).

The situations studied in this dissertation still represent "small worlds" (Binmore, 2008; Savage, 1954): People were presented with the state of the world, possible consequences and alternatives relevant to their judgment and decision, even if these were presented as uncertain. This approach may allow the decision processes to be studied, but it sidesteps an important and challenging step preceding judgment and decision making in large world moral situations: awareness of and search for relevant information.

Can we develop computational models of moral judgments and decisions? The ecological rationality perspective holds the promise of fostering a better understanding of judgment and decision making in moral situations by attending to the match between heu-

ristics processes and the environment. However, this requires the development of computational models which clearly specify the information that is taken into account and how it is processed. While some attempts have been made (e.g., Coenen & Marewski, 2009; Mikhail, 2007), most of the models proposed to date either do not specify the conditions that bring about one judgment instead of another (e.g., Haidt, 2001) or subscribe to some kind of rather broad and vague dual-system model (Greene et al., 2001).

However, developing computational models for the moral domain is challenging. Different people rely on different processes. This heterogeneity in moral judgment and decision making requires researchers to classify people accordingly instead of relying on aggregate results (Fischbacher, Hertwig, & Bruhin, 2013). At least for moral judgments, this may well require the research paradigm to be changed to cover a larger number of similar judgment situations. Yet given the context sensitivity of moral judgments, it is far from clear whether this is possible at all without trivially limiting the scope. In addition, moral judgments are affected by a variety of further factors and related to emotions in a close but yet to be understood way (for a review, see Waldmann, Nagel, & Wiegmann, 2012).

To some extent, the situation appears to be better within the study of economic games, which provide more neutral situations than the vignettes used in research on moral judgment. Yet this neutrality comes at the price of excluding potentially important cues that normally would be available in situations outside the lab. And while some researchers have worked on agent-based process models (e.g., Fischbacher et al., 2013; Gonzalez & Martin, 2011; Roth & Erev, 1995), the focus is rather on social preference models, which again bet on stable psychometric parameters instead of investigating cognitive mechanisms and their interaction with the environment.

How can we develop a systematic theory of the structure of social environments? Actually predicting moral judgment and behavior not only requires an analysis of the informational constraints stemming from the natural environment. It also requires a systematic theory of the structure of social environments in order to predict which heuristics will (or can) be used in a particular situation. Important constraints may be, for instance, the structure of social interconnections, which contains information about the position of a person within their social network, their interaction partners and information

sources; the stability of the social environment as resulting from norms and institutionalization; and the goals and behavior of others we interact with.

One starting point to understand which heuristics may be used in which situations is to identify across cultures some basic domains of morality characterized by particular problems, stimuli, and emotions, such as Shweder's *three ethics* (Shweder, Much, Mahaptra, & Park, 1997) or Haidt's *five foundations* (Haidt & Joseph, 2007). Another approach is provided by Fiske's taxonomy of four kinds of relationships (Fiske, 1992; Rai & Fiske, 2011)—equality matching, market pricing, authority ranking, and community sharing relations—which constrain the set of heuristics that may be selected within a given situation.

Final Thoughts

The research presented in this dissertation contributes to the study of judgment and decision making in moral situations by analyzing it from an ecological rationality perspective and thus incorporating different degrees of uncertainty within the empirical study of two widely used research paradigms. A better understanding of judgment and decision making in moral situation is important not only to understand when and why humans succeed or fail in pursuing certain ethical goals, but also to design better environments capable of promoting the values upon which a society agrees.

References

Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*, 737–753.

Anscombe, G. E. M. (1958). Modern moral philosophy. *Philosophy*, *33*, 1–19. doi:10.1017/S0031819100037943

Aquinas, T. (1981). *The Summa Theologica of St. Thomas Aquinas*. Notre Dame: Christian Classics.

Arendt, H. (1964). *Eichmann in Jerusalem. A Report on the banality of evil.* New York: Penguin.

Aristotle (1984). Nichomachean Ethics. In J. Barnes (Ed.), *Complete works of Aristotle* (Vol. 2, pp. 1729–1867). Princeton: Princeton University Press.

Axelrod, R. (1986). An evolutionary approach to norms. *The American Political Science Review*, *80*, 1095–1111.

Baron, J. (1994). Nonconsequentialist decisions. *Behavioral and Brain Sciences*, *17*, 1–10. doi:10.1017/S0140525X0003301X

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, *54*, 569–579. doi:10.1037/0022-3514.54.4.569

Baron, J., & Ritov, I. (2009). Protected values and omission bias as deontological judgments. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 50, pp. 133–167). San Diego: Academic Press.

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *The Academy of Management Review*, *3*, 439–449.

Bennis, W. M., Medin, D. L., & Bartels, D. M. (2010). The costs and benefits of calculation and moral rules. *Perspectives on Psychological Science*, *5*, 187–202. doi:10.1177/1745691610362354

Bereby-Meyer, Y., & Roth, A. E. (2006). The speed of learning in noisy games: Partial reinforcement and the sustainability of cooperation. *American Economic Review*, *96*, 1029–1042. doi:10.1257/aer.96.4.1029

Berg, N., & Gigerenzer, G. (2010). As-if behavioral economics: Neoclassical economics in disguise? *History of Economic Ideas*, *18*, 133-165. doi:10.1400/140334

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, *20*, 351–368. doi:10.1093/pan/mpr057

Binmore, K. (2008). *Rational Decisions*. Princeton: Princeton University Press.

Binmore, K., & Shaked, A. (2010). Experimental economics: Where next? *Journal of Economic Behavior & Organization*, *73*, 87–100. doi:10.1016/j.jebo.2008.10.019

Birnbaum, M. H., Coffey, G., Mellers, B. A., & Weiss, R. (1992). Utility measurement: Configural-weight theory and the judge's point of view. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 331–346.

Blanco, M., Engelmann, D., & Normann, H. T. (2010). A within-subject analysis of other-regarding preferences. *Games and Economic Behavior, 72,* 321–338.

Blasi, A. (2009). The moral functioning of mature adults and the possibility of fair moral reasoning. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, identity and character. Explorations in Moral Psychology* (pp. 396–440). New York: Cambridge University Press.

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin*, *88*, 1–45. doi:10.1037/0033-2909.88.1.1

Blass, T. (1991). Understanding behavior in the Milgram obedience experiment: The role of personality, situations, and their interactions. *Journal of Personality and Social Psychology*, *60*, 398–413.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, *90*, 166–193.

Boyd, R., & Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 3281–3288. doi:10.1098/rstb.2009.0134

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: making choices without trade-offs. *Psychological Review*, *113*, 409–432. doi:10.1037/0033-295X.113.2.409

Brighton, H., & Gigerenzer, G. (2012). Are rational actor models "rational" outside small worlds? In K. Binmore & S. Okasha (Eds.), *Evolution and Rationality*. Cambridge University Press. Retrieved from http://dx.doi.org/10.1017/CBO9780511792601.006

Bröder, A. (2012). The quest for take-the-best: Insights and outlooks from experimental research. In P. M. Todd, G. Gigerenzer, & ABC Research Group (Eds.), *Ecological Rationality: Intelligence in the World,* (pp.216-240). New York: Oxford University Press.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? *Perspectives on Psychological Science*, *6*, 3–5. doi:10.1177/1745691610393980

Burger, J. M. (2009). Replicating Milgram: Would people still obey today? *American Psychologist*, *64*, 1–11. doi:10.1037/a0010932

Camerer, C. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton: Princeton University Press..

Christensen, J. F., & Gomila, A. (2012). Moral dilemmas in cognitive neuroscience of moral decision-making: A principled review. *Neuroscience & Biobehavioral Reviews*, *36*, 1249–1264. doi:10.1016/j.neubiorev.2012.02.008

Christensen-Szalanski, J. J.., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *48*, 147–168. doi:10.1016/0749-5978(91)90010-Q

Chugh, D, Bazerman, M. H., & Banaji, M. R. (2005). Bounded ethicality as a psychological barrier to recognizing conflicts of interest. In D. Moore, D. Cain, G. Loewenstein, & M. Bazerman (Eds.), *Conflicts of Interest: Problems and Solutions from Law, Medicine and Organizational Settings.* London: Cambridge University Press.

Chugh, D., & Bazerman, M. (2007). Bounded awareness: What you fail to see can hurt you. *Mind and Society: Cognitive Studies in Economics and Social Sciences*, *6*, 1–18.

Coenen, A., & Marewski, J. N. (2009). Predicting moral judgments of corporate responsibility with formal decision heuristics. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1524–1528). Austin: Cognitive Science Society.

Coombs, C., Dawes, R., & Tversky, A. (1970). *Mathematical Psychology: An Elementary Introduction*. Oxford: Prentice-Hall.

Cosmides, L., & Tooby, J. (2004). Knowing thyself: The evolutionary psychology of moral reasoning and moral sentiments. In R. E. Freeman & P. Werhane (Eds.), *Business, Science, and Ethics* (pp. 91–127). Charlottesville, VA: Society for Business Ethics.

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*, 353–380. doi:10.1016/j.cognition.2008.03.006

Daniels, D., Dunn, J., Furstenberg, F. F., & Plomin, R. (1985). Environmental differences within the family and adjustment differences within pairs of adolescent siblings. *Child Development*, *56*, 764–774.

Darley, J. M., & Batson, C. D. (1973). "From Jerusalem to Jericho": A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, *27*, 100–108. doi:10.1037/h0034449

Darwin, C. (1871). *The Descent of Man*. Princeton, NJ: Princeton University Press.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571–582. doi:10.1037/0003-066X.34.7.571

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, *22*, 1915–1953.

DeScioli, P., & Kurzban, R. (2009). Mysteries of morality. *Cognition*, *112*, 281–299. doi:10.1016/j.cognition.2009.05.008

Dhami, M. K. (2003). Psychological models of professional decision making. *Psychological Science*, *14*, 175–180. doi:10.1111/1467-9280.01438

Dhami, M. K., & Ayton, P. (2001). Bailing and jailing the fast and frugal way. *Journal of Behavioral Decision Making*, *14*, 141–168. doi:10.1002/bdm.371

Doris, J. M. (2002). *Lack of Character: Personality and Moral Behavior*. Cambridge University Press.

Einhorn, H. J., & Hogarth, R. M. (1975). Unit weighting schemes for decision making. *Organizational Behavior and Human Performance*, *13*, 171–192. doi:10.1016/0030-5073(75)90044-6

Elms, A., & Milgram, S. (1966). Personality characteristics associated with obedience and defiance toward authoritative command. *Journal of Experimental Research in Personality*, *1*, 282–289.

Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, *37*, 1097–1126. doi:10.1037/0022-3514.37.7.1097

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*, 817–868.

Feldman, R. (2009). *The Liar in Your Life: The Way to Truthful Relationships*. New York: Twelve.

Feltz, A., & Cokely, E. T. (2009). Do judgments about freedom and responsibility depend on who you are? Personality differences in intuitions about compatibilism and incompatibilism. *Consciousness and Cognition*, *18*, 342–350. doi:10.1016/j.concog.2008.08.001

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*, 171–178. doi:10.1007/s10683-006-9159-4

Fischbacher, U., Hertwig, R., & Bruhin, A. (2013). How to model heterogeneity in costly punishment: Insights from responders' response times. *Journal of Behavioral Decision Making*. doi: 10.1002/bdm.1779

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299. doi:10.1037/0096-1523.1.3.288

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, *99*, 689–723.

Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, *5*, 5–15.

Foot, P. (1978). *Virtues and vices and other essays in moral philosophy*. Berkeley: : University of California Press.

Friedman, M. (1953). *Essays in Positive Economics*. Chicago: University of Chicago Press.

Friedman, M. (1976). *Price Theory*. Chicago: Aldine De Gruyter.

Gangadharan, L., & Nemes, V. (2009). Experimental analysis of risk and uncertainty in provisioning private and public goods. *Economic Inquiry*, *47*, 146–164. doi:10.1111/j.1465-7295.2007.00118.x

Gigerenzer, G. (2008). *Rationality for Mortals*. New York: Oxford University Press.

Gigerenzer, G. (2010). Moral satisficing: Rethinking moral behavior as bounded rationality. *Topics in Cognitive Science*, *2*, 528–554. doi:10.1111/j.1756-8765.2010.01094.x

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, *1*, 107–143. doi:10.1111/j.1756-8765.2008.01006.x

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451–482. doi:10.1146/annurev-psych-120709-145346

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669. doi:10.1037/0033-295X.103.4.650

Gigerenzer, G., & Regier, T. (1996). How do we tell an association from a rule? Comment on Sloman (1996). *Psychological Bulletin*, *119*, 23–26. doi:10.1037//0033-2909.119.1.23

Gigerenzer, G., & Selten, R. (2001). *Bounded rationality:The adaptive toolbox*. Cambridge, MA: MIT Press.

Gigerenzer, G., & Sturm, T. (2012). How (far) can rationality be naturalized? *Synthese*, *187*, 243–268. doi:10.1007/s11229-011-0030-6

Gigerenzer, G., Todd, P. M., & The ABC Research Group. (1999). *Simple Heuristics That Make Us Smart*. Oxford: Oxford University Press.

Gino, F., Shu, L., & Bazerman, M. (2008). Nameless Harmless = Blameless: When Seemingly Irrelevant Factors Influence Judgment of (Un)ethical Behavior. *Program On Negotiation (PON) at Harvard Law School Working Paper Series*. doi:10.2139/ssrn.1238661

Goeree J. K., Holt C. A., & Palfrey T. R. (2003). Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior*, *45*, 97–113. doi:10.1016/S0899-8256(03)00052-6

Gonzalez, C., & Martin, J. M. (2011). Scaling up instance-based learning theory to account for social interactions. *Negotiation and Conflict Management Research*, *4*, 110–128. doi:10.1111/j.1750-4716.2011.00075.x

Gourevitch, P. (1995, December 18). After the genocide. When a people murders up to a million fellow-countrymen, what does it mean to survive? *The New Yorker*. Retrieved from

http://www.newyorker.com/archive/1995/12/18/1995_12_18_078_TNY_CARDS_000372942

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, *111*, 364–371. doi:10.1016/j.cognition.2009.02.001

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*, 2105–2108.

Greene, J. D., Morelli, S., Lowenberg, K., Nystrom, L., & Cohen, J. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*, 1144–1154. doi:10.1016/j.cognition.2007.11.004

Greiner, B. (2004). *The online recruitment system ORSEE 2.0 - A guide for the organization of experiments in economics* (Working Paper Series in Economics No. 10). University of Cologne, Department of Economics. Retrieved from http://ideas.repec.org/p/kls/series/0010.html

Guilbault, R. L., Bryant, F. B., Brockway, J. H., & Posavac, E. J. (2004). A meta-analysis of research on hindsight bias. *Basic and Applied Social Psychology*, *26*, 103–117. doi:10.1080/01973533.2004.9646399

Güth, W., Levati, M. V., Sutter, M., & Van der Heijden, E. (2007). Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics*, *91*, 1023–1042. doi:10.1016/j.jpubeco.2006.10.007

Guyer, M. J., & Rapoport, A. (1972). 2×2 games played once. *The Journal of Conflict Resolution*, *16*, 409–431.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, *108*, 814–834.

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*, 998–1002. doi:10.1126/science.1137651

Haidt, J., & Bjorklund, F. (2008). Social intuitionists answer six questions about moral psychology. In W. Sinnott-Armstrong (Ed.), *Moral Psychology, Volume 2: The Cognitive Science of Morality: Intuition and Diversity: Cognitive Science of Morality: Intuition and Diversity v. 2* (pp. 181–217). Cambridge, MA: MIT Press.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*, 55–66.

Haidt, J., & Joseph, C. (2007). The moral mind: How 5 sets of innate moral intuitions guide the development of many culture-specific virtues, and perhaps even modules. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The Innate Mind: Foundations and the Future* (Vol. 3, pp. 367–392). Oxford: Oxford University Press, USA.

Haidt, J., & Kesebir, S. (2010). Morality. In S. Fiske & D. Gilbert (Eds.), *Handbook of Social Psychology*. New Jersey: John Wiley & Sons.

Hardin, G. (1968). The tragedy of the commons. *Science, 162 (3859),* 1243-1248. *doi:10.1126/science.162.3859.1243*

Harless, D. W., & Camerer, C. F. (1994). The predictive utility of generalized expected utility theories. *Econometrica*, *62*, 1251–1289. doi:10.2307/2951749

Harman, G. (1999). Moral philosophy meets social psychology: Virtue ethics and the fundamental attribution error. *Proceedings of the Aristotelian Society*, *99*, 315–331.

Hau, R., Pleskac, T. J., & Hertwig, R. (2010). Decisions from experience and statistical probabilities: Why they trigger different choices than a priori probabilities. *Journal of Behavioral Decision Making*, *23*, 48–68. doi:10.1002/bdm.665

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description-experience gap in risky choice: the role of sample size and experienced probabilities. *Journal of Behavioral Decision Making*, *21*, 493–518. doi:10.1002/bdm.598

Hauser, M. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. Ecco.

Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind & Language*, *22*, 1–21. doi:10.1111/j.1468-0017.2006.00297.x

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, *107*, 311–327. doi:10.1037/0033-2909.107.3.311

Herrera-Estrella, L., & Alvarez-Morales, A. (2001). Genetically modified crops: hope for developing countries? *EMBO reports*, *2*, 256–258. doi:10.1093/embo-reports/kve075

Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, *317*, 1360–1366. doi:10.1126/science.1146282

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*, 534 –539. doi:10.1111/j.0956-7976.2004.00715.x

Hertwig, R., Davis, J. N., & Sulloway, F. J. (2002). Parental investment: How an equity motive can produce inequality. *Psychological Bulletin*, *128*, 728–745.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, *13*, 517–523. doi:10.1016/j.tics.2009.09.004

Hertwig, R., & Herzog, S. M. (2009). Fast and frugal Heuristics: Tools of social rationality. *Social Cognition*, *27*, 661–698. doi:10.1521/soco.2009.27.5.661

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, *62*, 1291–1326. doi:10.2307/2951750

Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge updating? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 566–581. doi:10.1037/0278-7393.26.3.566

Kunreuther, H., Silvasi, G., Bradlow, E. T., & Small, D. (2009). Bayesian analysis of deterministic and stochastic prisoner's dilemma games. *Judgment and Decision Making*, *4*, 363–384.

Human Rights Watch. (2002). *Saudi Arabia: Religious police role in school fire criticized* (News). Human Rights Watch. Retrieved from http://www.hrw.org/en/news/2002/03/14/saudi-arabia-religious-police-role-school-fire-criticized

Hume, D. (1983). *An enquiry concerning the principles of morals*. Indianapolis: Hackett Publishing.

Iturbe-Ormaetxe, I., Ponti, G., Tomás, J., & Ubeda, L. (2011). Framing effects in public goods: Prospect theory and experimental evidence. *Games and Economic Behavior*, *72*, 439–447.

Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, *7*, 35–51.

Johnson, E. J., & Goldstein, D. G. (2003). Do defaults save lives? *Science*, *302*, 1338–1339.

Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue-contingent model. *The Academy of Management Review*, *16*, 366–395.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 292, 263.

Katz, F. E. (1993). *Ordinary people and extraordinary evil: A report on the beguilings of evil*. Albany: State University of New York Press.

Kelly, D., Stich, S., Haley, K. J., Eng, S. J., & Fessler, D. M. T. (2007). Harm, affect, and the moral/conventional distinction. *Mind & Language*, *22*, 117–131. doi:10.1111/j.1468-0017.2007.00302.x

Knight, F. H. (1921). *Risk, Uncertainty, and Profit*. Boston, MA: Houghton Mifflin.

Knobe, J., & Nichols, S. (2007). An experimental philosophy manifesto. In J. Knobe (Ed.), *Experimental Philosophy*. Oxford: Oxford University Press.

Kohlberg, Lawrence. (1984). *The psychology of moral development: Essays on moral development.* (Vol. 2). New York: Harper & Row.

Kohlberg, L., & Candee, D. (1978). The relationship of moral judgment to moral action. In Lawrence Kohlberg (Ed.), *Essays in moral development* (Vol. 2, pp. 498–581). New York: Harper & Row.

Kohlberg, L., Levine, C., & Hewer, A. (1983). *Moral stages: a current formulation and a response to critics*. New York: Karger.

Krebs, D. L. (2008). Morality: An evolutionary account. *Perspectives on Psychological Science*, *3*, 149–172. doi:10.1111/j.1745-6924.2008.00072.x

Krebs, D. L., & Denton, K. (2005). Toward a more pragmatic approach to morality: A critical evaluation of Kohlberg's model. *Psychological Review*, *112*, 629–649.

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, *118*, 97–109. doi:10.1037/a0020762

Langlois, R. N., & Cosgel, M. M. (1993). Frank Knight on risk, uncertainty, and the firm: A new interpretation. *Economic Inquiry*, *31*, 456–465. doi:10.1111/j.1465-7295.1993.tb01305.x

Lapsley, D. K., & Hill, P. L. (2008). On dual processing and heuristic approaches to moral cognition. *Journal of Moral Education*, *37*, 313. doi:10.1080/03057240802227486

Latané, B., & Rodin, J. (1969). A lady in distress: Inhibiting effects of friends and strangers on bystander intervention. *Journal of Experimental Social Psychology*, *5*, 189–202. doi:10.1016/0022-1031(69)90046-8

Leland, J. W. (1994). Generalized similarity judgments: An alternative explanation for choice anomalies. *Journal of Risk and Uncertainty*, *9*, 151–172. doi:10.1007/BF01064183

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*, 255–275.

LeRoy, S. F., & Singell, L. D. (1987). Knight on risk and uncertainty. *Journal of Political Economy*, *95*, 394–406. doi:10.2307/1832078

Levati, M. V., Morone, A., & Fiore, A. (2009). Voluntary contributions with imperfect information: An experimental study. *Public Choice*, *138*, 199–216. doi:10.1007/s11127-008-9346-2

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, *21*, 153–174. doi:10.1257/jep.21.2.153

Lo, A. W., & Mueller, M. T. (2010). WARNING: Physics Envy May Be Hazardous To Your Wealth. *Journal of Investment Management*, *8,* 13–63.

Ludvig, E. A., & Spetch, M. L. (2011). Of black swans and tossed coins: Is the description-experience gap in risky choice limited to rare events? *PLoS ONE*, *6*, e20262. doi:10.1371/journal.pone.0020262

MacIntyre, A. (1981). *After Virtue: A Study in Moral Theory.* Notre Dame, IN: University of Notre Dame Press.

Marewski J. N. & Krol K. (2010). Fast, frugal, & moral: Uncovering the heuristics of morality. *Journal of Organizational Moral Psychology*, *1*, 1-20.

McCarter, M. W., Rockmann, K. W., & Northcraft, G. B. (2010a). Is it even worth it? The effect of loss prospects in the outcome distribution of a public goods dilemma. *Organizational Behavior and Human Decision Processes*, *111*, 1–12. doi:10.1016/j.obhdp.2009.06.003

Messick, D. M., & Bazerman, M. H. (2001). Ethical leadership and the psychology of decision making. *Research in Ethical Issues in Organizations*, *3*, 213–238.

Messick, D. M., & Schell, T. (1992). Evidence for an equality heuristic in social decision making. *Acta Psychologica*, *80*, 311–323. doi:10.1016/0001-6918(92)90053-G

Mikhail, J. (2007). Universal moral grammar: theory, evidence and the future. *Trends in Cognitive Sciences*, *11*, 143–152. doi:10.1016/j.tics.2006.12.007

Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, *67*, 371–378. doi:10.1037/h0040525

Milgram, S. (1974). *Obedience to Authority: An Experimental View*. New York: Harper & Row.

Gong, M., Baron, J., & Kunreuther, H. (2009). Group cooperation under uncertainty. *Journal of Risk and Uncertainty*, *39*, 251–270.

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

Nisbett, R. E., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259.

Oakley, J. (1996). Varieties of virtue ethics. *Ratio*, *9*, 128–152. doi:10.1111/j.1467-9329.1996.tb00101.x

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

Palazzo, G., Krings, F., & Hoffrage, U. (2012). Ethical blindness. *Journal of Business Ethics*, *109*, 323–338. doi:10.1007/s10551-011-1130-4

Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). *Running Experiments on Amazon Mechanical Turk* (SSRN Scholarly Paper No. ID 1626226). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=1626226

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.

Persijn, G. G., & Van Netten, A. R. (1997). Public education and organ donation. *Transplantation Proceedings*, *29*, 1614–1617.

Pettit, D., & Knobe, J. (2009). The pervasive impact of moral judgment. *Mind & Language*, *24*, 586–604. doi:10.1111/j.1468-0017.2009.01375.x

Pichert, D., & Katsikopoulos, K. V. (2008). Green defaults: Information presentation and pro-environmental behaviour. *Journal of Environmental Psychology*, *28*, 63–73.

Prinz, J. (2009). The normativity challenge: Cultural psychology provides the real threat to virtue ethics. *The Journal of Ethics*, *13*, 117–144. doi:10.1007/s10892-009-9053-3

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, *118*, 57–75. doi:10.1037/a0021867

Rakow, T., Demes, K. A., & Newell, B. R. (2008). Biased samples not mode of presentation: Re-examining the apparent underweighting of rare events in experience-based choice. *Organizational Behavior and Human Decision Processes*, *106*, 168–179. doi:10.1016/j.obhdp.2008.02.001

Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making*, *23*, 1–14.

Rest, J. R. (1986). *Moral development: Advances in Research and Theory*. New York: Praeger Publishers.

Ridley, M. (1996). *The Origins of Virtue: Human Instincts and the Evolution of Cooperation* . New York: Penguin Books.

Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics and how can we tell. In Gerd Gigerenzer, P. M. Todd, & ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 141-167). Oxford: Oxford University Press.

Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, *127*, 258–276. doi:10.1016/j.actpsy.2007.05.004

Rieskamp, J., & Otto, P. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*, 236, 207.

Rithalia, A., McDaid, C., Suekarran, S., Myers, L., & Sowden, A. (2009). Impact of presumed consent for organ donation on donation rates: A systematic review. *BMJ*, *338*, a3162. doi:10.1136/bmj.a3162

Ritov, I., & Baron, J. (1999). Protected values and omission bias. *Organizational Behavior and Human Decision Processes*, *79*, 79–94. doi:10.1006/obhd.1999.2839

Rosenbaum, J. E. (2009). Patient teenagers? A comparison of the sexual behavior of virginity pledgers and matched nonpledgers. *Pediatrics*, *123*, e110–e120. doi:10.1542/peds.2008-0407

Ross, L., & Nisbett, R. E. (1991). *The Person and the Situation: Perspectives of Social Psychology*. New York: McGraw-Hill.

Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, *8*, 164–212. doi:10.1016/S0899-8256(05)80020-X

Rozin, P. (1999). The process of moralization. *Psychological Science*, *10*, 218–221. doi:10.1111/1467-9280.00139

Rubinstein, A. (1988). Similarity and decision-making under risk (is there a utility theory resolution to the Allais paradox?). *Journal of Economic Theory*, *46*, 145–153. doi:10.1016/0022-0531(88)90154-8

Savage, L. J. (1951). The Theory of Statistical Decision. *Journal of the American Statistical Association*, *46*, 55–67. doi:10.2307/2280094

Savage, L. J. (1954). *The Foundations of Statistics*. New York: John Wiley and Sons.

Scheffler, S. (1988). *Consequentialism and Its Critics*. Oxford, New York: Oxford University Press.

Schooler, C. (1972). Birth order effects: not here, not now. *Psychological Bulletin*, *78*, 161–175.

Shweder, R. A., Much, N. C., Mahaptra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, and divinity), and the "big three" explanations of suffering, as well. In A. Brandt & P. Rozin (Eds.), *Morality and health* (pp. 119–169). New York: Routledge.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*, 99–118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129–138. doi:10.1037/h0042769

Simon, H. A. (1979). Rational decision making in business organizations. *The American Economic Review*, *69*, 493–513. doi:10.2307/1808698

Simon, H. A. (1990). A mechanism for social selection and successful altruism. *Science*, *250*, 1665–1668.

Smith, V. L. (2003). Constructivist and ecological rationality in economics. *American Economic Review*, *93*, 465–508.

Smith, V. L. (2008). *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge: Cambridge Cambridge University Press.

Sollie, P. (2009). On uncertainty in ethics and technology. In P. Sollie & M. Düwell (Eds.), *Evaluating New Technologies* (pp. 141–158). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-90-481-2229-5_10

Sprouse, J. (2010). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, *43*, 155–167. doi:10.3758/s13428-010-0039-7

Sunstein, C. R. (2003). Terrorism and probability neglect. *Journal of Risk and Uncertainty*, *26*, 121–136.

Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences*, *28*, 531–542. doi:10.1017/S0140525X05000099

Tenbrunsel, A. E., & Smith-Crowe, K. (2008). Ethical decision making: Where we've been and where we're going. *The Academy of Management Annals*, *2*, 545. doi:10.1080/19416520802211677

Todd, P. M., Gigerenzer, G., & The ABC Research Group. (2012). *Ecological Rationality: Intelligence in the World*. New York: Oxford University Press.

Trevino, L. K. (1986). Ethical decision making in organizations: A person-situation interactionist model. *The Academy of Management Review*, *11*, 601–617.

Turiel, E. (1983). *The Development of Social Knowledge: Morality and Convention*. Cambridge: Cambridge University Press.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*, 281–299. doi:10.1037/h0032955

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323. doi:10.1007/BF00122574

Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychological Science*, *20*, 473 –479. doi:10.1111/j.1467-9280.2009.02319.x

Unger, P. K. (1996). *Living High and Letting Die: Our Illusion of Innocence*. Oxford: Oxford University Press.

UNHCR. (2010). *Number of forcibly displaced rises to 43.3 million last year, the highest level since mid-1990s* (News Stories). UNHCR. Retrieved from http://www.unhcr.org/4c176c969.html

Vollan, B., & Ostrom, E. (2010). Cooperation and the commons. *Science, 330(6006),* 923-924. doi:10.1126/science.1198349

Von Dijk, E., Wit, A., Wilke, H., & Budescu, D. V. (2004). What we know (and do not know) about the effects of uncertainty on behavior in social dilemmas. In R. Suleiman, D. V. Budescu, I. Fischer, & D. M. Messick (Eds.), *Contemporary psychological research on social dilemmas* (pp. 315–331). Cambridge: Cambridge University Press.

Wakker, P. P. (2010). *Prospect Theory: For risk and ambiguity*. Cambridge: Cambridge University Press.

Wald, A. (1945). Statistical decision functions which minimize the maximum risk. *The Annals of Mathematics*, *46*, 265. doi:10.2307/1969022

Waldmann, M. R., Nagel, J., & Wiegmann, A. (2012). Moral judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 364–389). Oxford: Oxford University Press.

Walster, E. (1967). "Second guessing" important events. *Human Relations*, *20*, 239–249. doi:10.1177/001872676702000302

Wiegmann, A., Okan, Y., & Nagel, J. (2012). Order effects in moral judgment. *Philosophical Psychology*. Advance online publication. doi: 10.1080/09515089.2011.631995

Wit, A., Van Dijk, E., Wilke, H., & Groenenboom, A. (2004). The interplay between environmental and social uncertainty in social dilemmas. In R. Suleiman, D. V. Budescu, I. Fischer, & Messick (Eds.), *Contemporary psychological research on social dilemmas* (pp. 376–398). Cambridge: Cambridge University Press.

World Disasters Report (2011). *Scribd*. Retrieved from http://www.scribd.com/doc/65875175/World-Disasters-Report-2011(January 24, 2013)

Zimbardo, P. G. (2008). *The Lucifer effect: Understanding how good people turn evil*. New York: Random House.

Zimmerman, M. J. (2010). *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge: Cambridge University Press.

Appendix A

Further Statistics and Experimental Materials for Chapter 3

A1. Further statistics

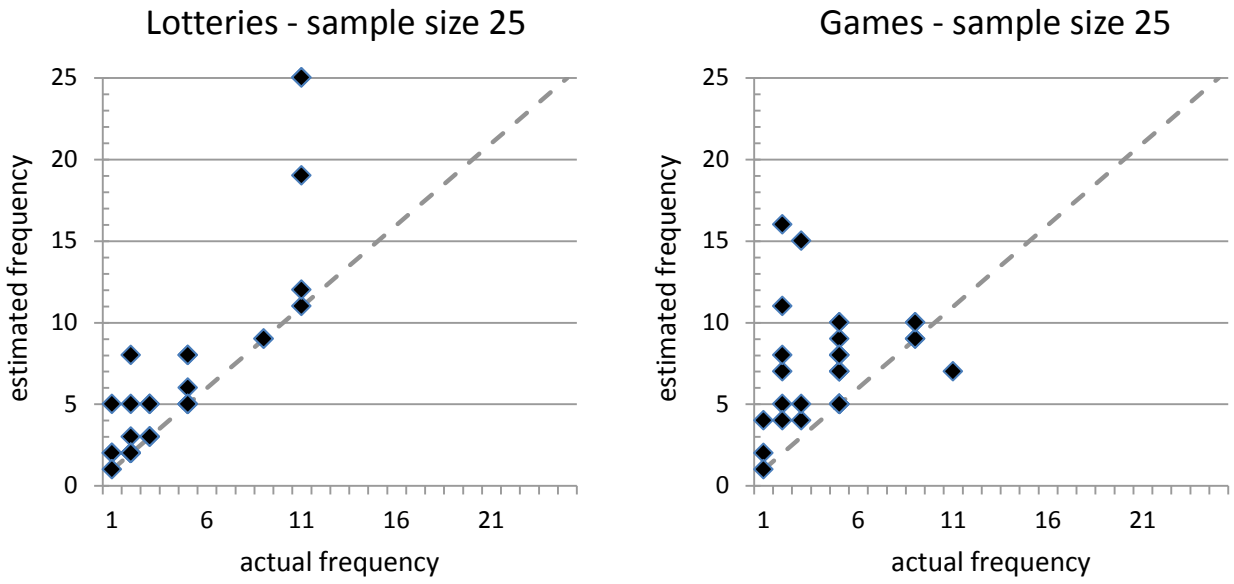


Figure A1. Participants' frequency estimates for the less likely event in the last round for lotteries and games.

Table A1.

Percentage of People Investing into the Risky Option if the Expected msr in the 2nd Half of the Drawn Sample is Larger than in the 1st Half of the Sample.

LOTTERY			
		Percentage of people (N=256)	
		not investing	investing
Expected outcome larger in 2nd half	no	20%	30%
	yes	14%	36%

STOCHASTIC PG			
		Percentage of people (N=256)	
		not investing	investing
Expected outcome larger in 2nd half	no	29%	18%
	yes	30%	22%

Table A2.

Percentage of Participants Who Selected Each Combination of Reasons for Contributing and not Contributing to the Public Good, in Description and Experience Conditions.

DESCRIPTION

		Contribution							
		Probability	Value	Conditional cooperation	Social uncertainty	Moral reason	Other reason	Never contributed	**Sum**
No contribution	Probability	41%	6%	6%	0%	3%	0%	0%	**56%**
	Value	0%	0%	0%	0%	0%	0%	0%	**0%**
	Conditional cooperation	0%	0%	3%	0%	3%	0%	0%	**6%**
	Social uncertainty	6%	0%	3%	0%	0%	0%	3%	**13%**
	Opportunism	13%	0%	0%	0%	0%	0%	3%	**16%**
	Other reason	0%	0%	3%	0%	3%	0%	0%	**6%**
	Always contributed	3%	0%	0%	0%	0%	0%	0%	**3%**
	Sum	**63%**	**6%**	**16%**	**0%**	**9%**	**0%**	**6%**	

EXPERIENCE

		Contribution							
		Probability	Value	Conditional cooperation	Social uncertainty	Moral reason	Other reason	Never contributed	**Sum**
No contribution	Probability	19%	6%	19%	0%	3%	0%	0%	**47%**
	Value	6%	6%	6%	0%	0%	0%	0%	**19%**
	Conditional cooperation	0%	3%	0%	0%	0%	0%	3%	**6%**
	Social uncertainty	3%	0%	3%	0%	0%	0%	0%	**6%**
	Opportunism	3%	6%	0%	0%	3%	0%	3%	**16%**
	Other reason	0%	0%	3%	0%	0%	0%	0%	**3%**
	Always contributed	0%	0%	3%	0%	0%	0%	0%	**3%**
	Sum	**31%**	**22%**	**34%**	**0%**	**6%**	**0%**	**6%**	

Note. 'Probability' refers to probability being the most important reason, 'value' refers to the value of the multiplier, 'cond. cooperation' refers to whether participants expected the other to cooperate or not, 'social uncertainty' indicates whether being not sure about what the other drives behavior, 'opportunism' indicates whether the player expected the other to cooperate and wanted to use this for her on benefits, 'other reason' indicates that some other reason than the those stated drive the decision, and 'always invested' indicates when participants always invested or did not invest.

A2. Questionnaire asking for Participants' Most Important Reasons to Contribute and Noncontribute in the stochastic Public Goods.

Participants indicated their most important reason for not contributing to the public good by selecting one of the following options:

1) probability - "The probability to get a low value for A seemed to high to me, independent of whether the other contributes.
2) value - "The two values for A were not high enough for me, independent of whether the other contributes.
3) conditional non-cooperation - "I thought that the other participant would not contribute."
4) social uncertainty - "I did not know what the other would do. Therefore I did not contribute."
5) greed/opportunism - "I thought that the other participant would contribute and this way I could profit without contributing myself
6) "None of the reasons applies."
7) "I always contributed."

Participants indicated their most important reason for contributing to the public good by selecting one of the following options:

1) probability - "The probability to get the high value for A seemed high enough to me, independent of whether the other contributes.
2) value - "The two values for A were high enough for me, independent of whether the other contributes.
3) conditional cooperation - "I thought that the other participant would contribute. 4) social uncertainty – "I did not know what the other would do. Therefore I contributed."
5) moral obligation – "I thought to contribute to the project is the morally right thing to do independent of what the other does."
6) "None of the reason applies.
7) "I never contributed."

A3. Instructions

[Anything in square brackets was not shown to the participants but serves to lead the reader through the instructions. The instructions pertain to the experience condition. At any point where participants are asked to 'draw' this is replaced in such a way that participants are told that they are shown probability statements.]

Dear Participant,
welcome and thank you very much for participating in this experiment!

For arriving on time you receive a payoff of 2.50€. In addition to this, you can earn more money through your decisions in this experiment. Therefore, please read the following instructions carefully. In the experiment you will make several decisions. At the end of the experiment the instructor will determine one round of the experiment at random which will account for your payoff.

Your total payoff from this experiment consists of the amount that you gained in the determined round plus 2.50€ for arriving on time. You will receive your total payoff in cash at the end of this experiment. The payment will happen privately. None of the other participants will know the amount of your payoff.

During the experiment, all amounts will be given in ECU (Experimental Currency Units). **Thereby 1 ECU = 1 Euro.**

Please, stay calm during the experiment and switch off your mobile phones! It is important that you do not take any notes or talk to other participants during the experiment. Communication among participants will automatically lead to a termination of the experiment in which case no participant will receive a payoff from the experiment. If case you have any questions or comments during the experiment, please raise your hand. One of the instructors will then come to your place and answer your question.

[Instructions lotteries]

Instructions

This part of the experiment consists of 8 rounds.
At the beginning of each round, you will receive 10 ECU. Please decide in each round between 2 options:

Option A: **The 10 ECU will be multiplied by 1. You will keep 10 ECU.**

Option B: **The 10 ECU will be multiplied by one of two possible values „A", one which is high and one which is lower.** Depending on which value of A will be actually realized, the 10 ECU will be increased or decreased.

You will learn about the likelihood of A attaining the high or the lower value by testing how often each of these values will occur. Each time you press on the button „draw", one of the two values of A will be drawn at random from a card deck and displayed on the screen. The more often a value occurs, the more likely it is.

The program will ask you to draw 25 times. Afterwards you will have to decide for option A or option B. Then, another round will follow.

Please note: You will be informed about the outcome of your decision only at the end of the experiment.

Example

You receive 10 ECU for this round. You have to decide between 2 options:

Option A: You keep the 10 ECU.

Option B: The 10 ECU will be multiplied with A. In this round A equals **0.5 with 10 % or 1.7 with 90%.**

You decide for option B. You will receive either 10 x 0.5 = 5 ECU with 10% or 10 x 1.7 = 17 ECU with 90% probability.

Your final payoff

At the end of the experiment the instructor will randomly determine one round for your final payoff. Then, he will determine one of the two values for A according to its likelihood, thus determining which value of A will be actually realized. The payoff that you gained in this round through your decision will be paid out to you together with the 2.50€ for arriving on time, at the end of the experiment.

The Experiment will now begin. From now on, all decisions that you will make will be relevant for your payoff. Please, remain quiet and seated during the whole experiment and do not talk to each other until you have left the room.

In case you have questions, please raise your hand!

[Instructions Games]

Instructions

Formation of groups
This part of the experiment consists of 10 rounds. In each round you will be part of a dyad interacting with another participant. You will not receive any information on who the other person in your group is; neither during nor after the experiment. For each round new groups will be formed such that no participant will be ever matched with the same counterpart again.

Your task
At the beginning of each round, each participant will receive 10 ECU. Both group members decide independently if they want to contribute 10 ECU to a project. Your payoff depends on your own decision and on the decision of the other group member in the respective round.

Option A: You do not contribute to the project and keep the 10 ECU.

Option B: You contribute to the project. All amounts will be multiplied by a value A. This determines the payoff from the project.

The payoff from the project will be equally split among the two group members; i.e., each group member receives half of the payoff, independent of whether or not they contributed to the project.

Your part of the project payoff = $\frac{\text{A * (your contribution + contribution of the other group member)}}{2}$

Then, a new round begins, in which you will be matched with another participant.

Your payoff per round:
Your payoff in each round consists of two parts:

- The amount of ECU, which you do not contribute
- Your part of the project payoff

Your payoff per round = amount of ECU, which you do not contribute + your part of the project

Example 1a:
A has a value of 1.7. You and your counterpart both contribute 10 ECU. The sum of these amounts is 2 x 10 ECU = 20 ECU. The total project payoff is 20 ECU x 1.7 = 34 ECU.
Each of you receives an equal share of this payoff; i.e., 34 ECU / 2 = 17 ECU.
Your total payoff from this round and the total payoff of your counterpart are both 17 ECU.

Example 1b:
A has a value of 1.7. Only you contribute 10 ECU to the project, but your counterpart does not. The sum of these amounts is 10 ECU. The total project payoff is 10 ECU x 1.7 = 17 ECU.
Each of you receives the same part of this payoff; i.e., 17 ECU / 2 = 8.50 ECU.
Your total payoff from this round is 8.50 ECU.
The total payoff of your counterpart is 8.50 ECU from the project plus 10 ECU which he kept for himself = 18.50 ECU.
Amount of the project payoff per round

The amount of the project payoff depends on the value of A. In the last two rounds, A has a fixed value. In each round, there are two possible values for A, a high value and a lower value. Depending on the value that is realized for A, the project payoff will be higher or lower. You will learn about the likelihood of A attaining the high or the lower value by testing how often each of these values will occur.
Each time you press on the button „draw", one of the two values of A will be drawn at random from a card deck and displayed on the screen. The more often a value occurs, the more likely it is.
The program will ask you to draw 25 times. Afterwards you will have to decide whether or not to contribute your 10 ECU to the project. Then, another round will follow.

Please note: You will be informed about the outcome of your decision only at the end of the experiment.

Example 2:
You receive 10 ECU for this round. You and the other group member decide independently from each other, whether or not to contribute your 10 ECU to the project. The project payoff depends on the sum of your contributions and the value of the multiplier A. In this round, A can take one of two values with following probabilities:

0.8 with 6% or 1.9 with 94%

You and your counterpart both contribute 10 ECU; i.e., in total 20 ECU.
The project payoff either equals 20 x 0.8 = 16 ECU with 6% or 20 x 1.9 = 38 ECU with 94% probability.
Each of you receives an equal share of the project payoff; i.e., either 16 ECU/2 = **8 ECU with 6%** or 38 ECU / 2 = **19 ECU with 94%**

Your final payoff
At the end of the experiment the instructor will randomly determine one round for your final payoff. Then, he will determine one of the two values for A according to its likelihood, thus determining which value of A will be actually realized. The payoff that you gained in this round through your decision will be paid out to you together with the 2.50€ for arriving on time, at the end of the experiment. Please, answer the following comprehension questions, before the experiment starts. Thereby we want to assure that all participants understand the rules of the experiment fully and correctly.
Please, remain quiet and seated during the whole experiment and do not talk to each other until you have left the room. In case you have questions, please raise your hand!

Appendix B

Further Statistics and Experimental Materials for Chapter 4

B1. Further statistics

Table B1

Descriptive Statistics for Moral Judgments per Condition for Each Dilemma and Across All Dilemmas.

	Moral judgments											
Dilemma	FS				HSbad				HSgood			
	N	*M*	*Mdn*	*SD*	*N*	*M*	*Mdn*	*SD*	*N*	*M*	*Mdn*	*SD*
Influenza	35	2.83	3	0.81	40	2.63	3	0.83	35	3.20	3	0.67
Torture	39	2.62	3	1.00	39	2.56	2	0.98	40	2.85	3	0.82
Screening	37	3.24	3	0.71	34	3.03	3	0.98	39	3.79	4	0.40
GM corn	42	2.69	3	0.80	35	2.29	2	0.81	34	3.15	3	0.65
Dam	38	2.74	3	0.64	39	2.46	2	0.87	37	2.89	3	0.65
Credit	37	2.57	3	0.68	36	2.89	3	0.61	35	2.97	3	0.77
Across all	228	2.78	3	0.82	223	2.64	3	0.89	220	3.15	3	0.75

Table B2

Comparison of Moral Ratings Between Conditions Separately for Each Dilemma and Across All Dilemmas.

	Moral judgments: comparisons between conditions											
Dilemma	FS/HSbad				FS/HSgood				HSbad/HSgood			
	U	*z*	*p*	*r*	*U*	*z*	*p*	*r*	*U*	*z*	*p*	*r*
Influenza	618.50	-0.94	.35	-0.11	463.50	-1.90	.06	-0.23	446.50	-2.95	.001	-0.34
Torture	728.00	-0.33	.74	-0.04	688.50	-0.94	.35	-0.11	642.50	-1.41	.16	-0.16
Screening	574.00	-0.67	.50	-0.08	416.50	-3.66	.001	-0.42	369.00	-3.75	.001	-0.44
GM corn	522.50	-2.35	.02	-0.27	510.50	-2.40	.02	-0.28	269.50	-4.16	.001	-0.50
Dam	585.00	-1.73	.08	-0.20	641.00	-0.76	.45	-0.09	516.50	-2.28	.02	-0.26
Credit	510.00	-1.93	.05	-0.23	464.50	-2.24	.03	-0.26	582.50	-0.60	.55	-0.07
Across all	23055.00	-1.83	.07	0.09	19124.00	-4.70	.001	0.22	16874.50	-6.03	.001	0.29

Note. Two-tailed Mann–Whitney U-test. An effect size of $r = .1$ indicates a small effect, $r = .3$ a medium effect, and $r = .5$ a large effect.

Table B3

Descriptive Statistics for Probability Estimates per Condition for Each Dilemma and Across All Dilemmas.

	Probability estimates of negative side effects											
Dilemma	FS				HSbad				HSgood			
	N	*M*	*Mdn*	*SD*	*N*	*M*	*Mdn*	*SD*	*N*	*M*	*Mdn*	*SD*
Influenza	35	33.37	30.00	19.68	40	50.78	50.00	35.97	35	22.89	20.00	22.93
Torture	39	52.56	50.00	27.82	39	74.23	80.00	26.01	40	29.60	22.50	27.37
Screening	37	20.97	15.00	18.91	34	35.15	27.50	26.02	39	17.10	10.00	14.97
GM corn	42	41.83	50.00	22.96	35	61.14	65.00	27.15	34	26.26	20.00	19.40
Dam	38	54.58	55.00	25.77	39	81.38	90.00	19.92	37	31.11	40.00	24.30
Credit	37	70.27	75.00	15.98	36	62.36	60.00	16.85	35	44.89	50.00	19.90
Across all	228	45.72	50.00	27.27	223	61.35	65.00	30.15	220	28.49	25.00	23.52

Table B4

Comparison of Probability Estimates for Negative Side Effects Between Conditions Separately for Each Dilemma and Across All D*ilemmas.*

	Probability estimates: comparisons between conditions											
Dilemma	FS/HSbad				FS/HSgood				HSbad/HSgood			
	U	*z*	*p*	*r*	*U*	*z*	*p*	*r*	*U*	*z*	*p*	*r*
Influenza	528.50	-1.82	.07	-0.21	421.50	-2.25	.02	-0.27	372.00	-3.49	.001	-0.40
Torture	422.00	-3.39	.001	-0.38	425.50	-3.48	.001	-0.39	204.50	-5.66	.001	-0.64
Screening	423.00	-2.38	.02	-0.28	657.00	-0.67	.50	-0.08	374.00	-3.21	.001	-0.38
GM corn	452.50	-2.90	.001	-0.33	437.00	-2.91	.001	-0.33	195.00	-4.82	.001	-0.58
Dam	302.00	-4.50	.001	-0.51	371.00	-3.53	.001	-0.41	98.50	-6.51	.001	-0.75
Credit	485.00	-2.01	.04	-0.23	208.50	-4.97	.001	-0.59	335.50	-3.41	.001	-0.40
Across all	17746.50	-5.56	.001	-0.26	15936.50	-6.69	.001	-0.32	9966.50	-10.84	.001	-0.51

Note. Two-tailed Mann–Whitney U-test. An effect size of $r = .1$ indicates a small effect, $r = .3$ a medium effect, and $r = .5$ a large effect.

Table B5

Spearman Rank Correlation Between Moral Ratings and Probability Estimates by Conditions Separately for Each Dilemma and Across All Dilemmas.

Dilemma	FS	HSbad	HSgood	Across conditions
	Spearman correlation			
Influenza	.01	- .28	- .11	- .24
Torture	- .11	- .27	- .11	- .19
Screening	- .52	- .51	- .04	- .46
GM corn	- .59	- .60	- .51	- .66
Dam	- .09	- .38	- .13	- .32
Credit	.08	- .02	- .60	- .26
Across all	- .32	- .35	- .29	- .39

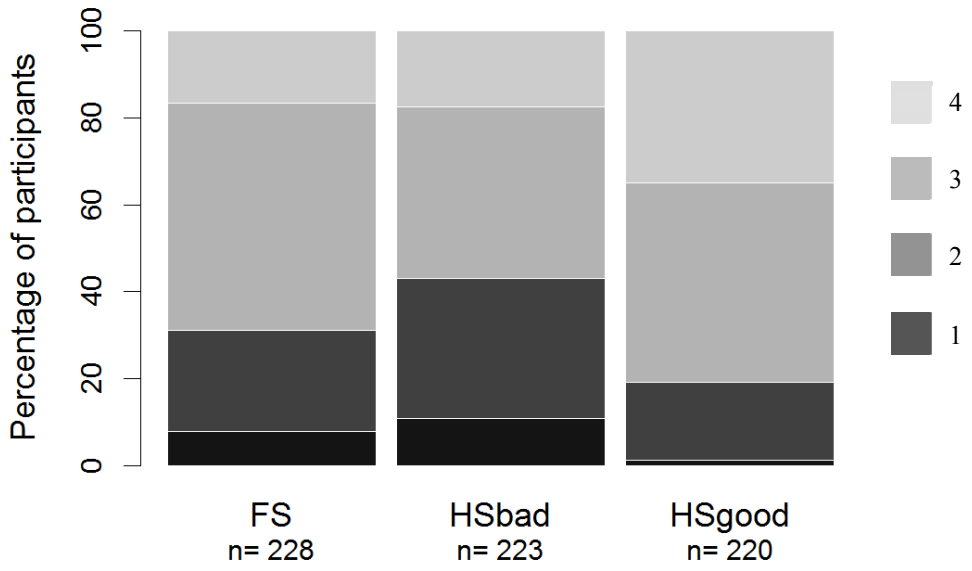


Figure B1. Percentage of participants on each level of the moral judgment scale (1-4) in *foresight* (FS), *hindsight bad* (HSbad), and *hindsight good* (HSgood) across dilemmas. The proportion of people rating the action to be completely or somewhat impermissible (scale rating 1 or 2) is highest in HSbad, lowest in HSgood, and in between in FS.

B2. The six moral dilemmas used in the study

Information shown in italics was additionally presented in the hindsight conditions (without italics). Notes in square brackets indicate whether this additional information was given in the hindsight bad or the hindsight good condition.

Influenza

The president of the World Health Organization has to make a decision in the following situation.

A new strain of flu has caused severe infections. 14 days after the first outbreak, about 30,000 infections have been reported on four continents. Further spread is thus inevitable. However, the severity of the influenza is hard to estimate. So far, about 1,000 deaths have been reported, yet the majority of patients experienced mild symptoms and fully recovered from the flu. Information from the southern hemisphere, which experienced the influenza wave earlier during their winter season, indicates that the number of deaths and severe infections were not higher than with the normal seasonal flu.

Initiating a large-scale vaccination campaign across countries could reduce severe infections and the number of deaths from the flu by up to 70%.

The vaccine for the virus has newly been developed. It can sometimes cause severe side effects that are as bad as the flu itself and potentially fatal. As the campaign would need to start immediately, further long-term tests are not possible. The vaccination will cost millions of dollars spent by governments around the world that would participate in the campaign, money that could be used elsewhere in healthcare.

In the end, the president decides to initiate a large-scale vaccination campaign.

The vaccination does not cause severe or even fatal side effects. [HSgood]
The vaccination causes severe and even fatal side effects in a number of cases. [HSbad]

Torture

A police officer has to make a decision in the following situation.

An eleven-year-old boy did not come home from school one day. Shortly thereafter, his parents, a wealthy banker's family, received a ransom note demanding 1 million dollars for the boy. After being notified by the parents, the police trailed the man who picked up the money, in the hopes that he would lead them to the boy. On the third day, the man still hadn't led them to the boy but instead booked a vacation and made preparations for the trip. The police then arrested the man who turned out to have no criminal record. During interrogations the man admitted that he was holding the boy captive and intimated that the boy was still alive. Yet he persisted in giving false information about his whereabouts. All independent search efforts by the police remained unsuccessful. The boy has now been missing for four days, a dangerously long time for a child without food and water.

The police officer in charge knows a way how he might get the information: by threatening the kidnapper that he will inflict pain on him if he doesn't tell where the boy is.

This violates the personal rights of the man who might as a consequence be acquitted due to the violation of procedural rules. Furthermore, it may damage the reputation of the police and lead to a debate about effective judicial constraints on police work.

In the end, the police officer decides to threaten the man that he will inflict pain on him.

At his trial, the kidnapper is convicted despite the violation of procedural rules. [HSgood]
At his trial, the kidnapper is being acquitted due to the violation of procedural rules. [HSbad]

Screening

A doctor has to make a decision about a recommendation in the following situation.

A woman of 50 asks whether to do a regular screening for detection of breast cancer which is one of the most prevalent cancers in women over 50. The screening with mammography uses X-ray to try to find breast cancer early when a cure is more likely.

The doctor informs the woman that a new systematic review estimates that screening leads only to a small reduction in the number of death from breast cancer.

Screening can also have disadvantages: a number of women without breast cancer will be falsely diagnosed and experience important psychological distress for many months. In addition, some healthy women will receive surgical treatment (partial or full removal of the breast) that is unnecessary because their cancers grow slowly and would not have led to death or sickness.

In the end, the doctor decides to recommend a regular screening.

Good: The screening does not lead to a wrong result and unnecessary treatment. [HSgood]
Bad: The screening leads to a wrong result and unnecessary treatment. [HSbad]

Genetically modified corn

The government of a developing country has to make a decision in the following situation.

Over the last decade, climate fluctuations repeatedly led to extreme weather conditions followed by crop failure and hunger crises. Many people died before emergency food aid from abroad could be provided. Extreme weather conditions would bring about a new famine, yet for the last two years the weather conditions have remained stable and the crop has been good.

A new genetically modified corn is available that could survive even under extreme weather conditions. A governmental program for its cultivation could prevent food scarcity. While there are other means to prevent food scarcity, all of them require long-term and structural changes and will not be effective within the next years.

The corn passed tests for food safety, but its long-term effects will only be known in the future. Due to the genetic modifications, it might possibly trigger allergies or diseases in humans and animals who consume it. Furthermore, it might reduce biological diversity and destroy the food chain for some animals. Due to these concerns, many industrial countries are still debating its use.

In the end, the government decides to start cultivating the modified corn.

The corn triggers no allergies and diseases in humans and the food chain of animal species remains unaffected. [HSgood]
The corn triggers allergies and diseases in a number of humans and destroys the food chain of some animal species. [HSbad]

Dam

The government of a developing country has to make a decision in the following situation.

40 out of 60 million people in the country live close to a large river that crosses from north to south and provides water resources for agriculture, which is the most important sector in the country. Over the last decade, severe floods repeatedly occurred that could not be restrained with standard flood protection. The consequences of the floods were devastating for the living conditions of farmers and the country's economy, yet in the last two years no further severe floods occurred.

Building a large dam could reduce the effects of severe floods by providing flood storage space and would also allow for perennial irrigation to improve agricultural conditions for the farmers.

The dam will flood an inhabited area upstream so that 2 million farmers will have to be relocated against their will to the cities, receiving only a one-time financial compensation for their land. In addition, ecological changes may result in lower water quality and higher costs of drinking water.

In the end, the government decides to build the dam.

The water quality remains stable and the costs of drinking water don't increase. [HSgood]
The water quality declines and the costs of drinking water increase. [HSbad]

Credit

The government has to make a decision in the following situation.

Due to the financial crisis, an increasing number of companies had to declare bankruptcy within the last 5 years. This year, a well-known traditional family business announced that it may soon be unable to remain in business because some of its major customers have not been paying their debts and the banks will not provide any credit. More than 5,000 jobs within the company and its suppliers will be lost if no new investors are found immediately.

The government considers supporting the company by providing an exceptional state credit that could help the company to survive their defaulting customers and to develop new markets.

To provide the credit, the government would need to reallocate funds that were intended to support creditworthy start-up businesses that do not get the loans they need from banks. As a result, a number of start-ups relying on state loans may not be able to survive and may have to declare insolvency, implying a loss of jobs.

In the end, the government decides to provide the company with the credit.

The start-ups manage to survive without state loans. [HSgood]
Many start-ups have to declare insolvency without state loans. [HSbad]

B3. Questionnaire asking for the most important reason

Think again about the moral judgment you made at the beginning. Which of the following best describes the most important reason underlying your moral judgment?

You judged the action to morally impermissible or somewhat morally impermissible, because:

- ☐ The harm of inaction was not sufficiently large / likely.
- ☐ The benefits of the action were not sufficiently large /likely.
- ☐ The negative consequences of the action were not sufficiently small / unlikely.
- ☐ The negative consequences of the action outweigh the benefits of the action.
- ☐ The action is morally wrong, and should not be permissible even if it leads to a large benefit or prevents a large harm in this case.
- ☐ Other reason:

You judged the action to be morally permissible or somewhat morally permissible, because:

- ☐ The harm of inaction was sufficiently large / likely.
- ☐ The benefits of the action were sufficiently large /likely.
- ☐ The negative consequences of the action were sufficiently small / unlikely.
- ☐ The benefits of the action outweigh the negative consequences of the action.
- ☐ The action is morally wrong, but should be permissible in this case because it leads to a large benefit or prevents a large harm.
- ☐ Other reason:

List of Tables

Chapter 3

Chapter 4

Appendix A

Appendix B

List of Figures

Chapter 2

Chapter 3

Chapter 4

Appendix A

Appendix B